Original Article
Identification and sequence analysis of aroA gene of avibacterium paragallinarum

Xue-Ze Lv, Hong-Jun Wang, Xiao-Ling Chen, Yu-Mei Gong, Yu-Rong Liang, Yun-Xia He, Pei-Jun Zhang

Beijing Key Laboratory for Prevention and Control of Infectious Diseases in Livestock and Poultry, Institute of Animal Husbandry and Veterinary Medicine, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

Abstract: Objectives: A convenient and efficient method for the clone of unknown full-length gene was established and the structure and molecule evolution of full-length aroA gene of Avibacterium paragallinarum were evaluated. Methods: This experiment took an international standard strain 145 (C-3) of Apg as template, designed degenerate primers according to aroA gene with CODEHOP software. The improved chromosome walking was used to amplify the whole gene sequence of aroA. Some structure researches on the nucleicacid and coding protein sequence have been done. Besides, we also did some research on molecular evolution analysis with other serotypes and related bacteria by sequence alignment. Results: Complete aroA gene of Avibacterium paragallinarum was obtained, which is 1 293 bp. The length of the gene encoded product is 430aa, which have three transmembrane domains and five potential epitopes. The homology of amino acid sequences is 91.4%-100% between different serotypes experiment strains, and the nucleic acid homology is over 74.2% between bacteria of the same family. Conclusions: This trial has combined the degenerate PCR and improved chromosome walking for the research of avibacterium coli aroA genes complete sequence for the first time, and obtained the expected results. It laid the foundation for the study of biological activity of its encoded protein and building avibacterium bacilli aroA deletion attenuated strains.

Keywords: Apg, aroA, gene identification, protein structure, molecule evolution

Introduction

Avibacterium paragallinarum (Apg) is a kind of short Gram-negative bacteria in pasteurellaceae [1], which can cause acute respiratory diseases in chickens. The disease has a worldwide distribution [2]. According to the typing methods of hemagglutination inhibition tests, Apg was divided into 3 serotypes of A (A-1, A-2, A-3, A-4), B (B-1) and C (C-1, C-2, C-3, C-4) [3]. Majority of inactivated vaccine currently used in the world generally were the A and C serotypes, but a large number of B serotype has recently been discovered in domestic [4]. Robert R.Bragg [5], a South Africa scholar, had confirmed that the C-3 serotype strains were highly virulent strains and it had spread outside of South Africa.

The aroA gene encodes a product of 5'-pyruvate shikimate-3-phosphate synthase (EPSPS), which is a very important enzyme in bacterial aromatic amino acid metabolic pathway. AroA

gene was firstly discovered in E.coli, then people have successfully cloned it from a variety of bacteria. At present, the research on the aroA gene of Avibacterium paragallinarum is limited to its partial sequence. Song et al. [6] has established a PCR diagnostic method for aroA genes of Avibacterium paragallinarum, which could amplify specific fragments of 800 bp in different serotypes of the bacterium. In this study, Avibacterium paragallinarum was taken as a template to establish a convenient and efficient method for the clone of unknown full-length genes, in order to enrich its molecular biology database and lay the foundation for future research.

Materials and methods

Experimental materials

Strains: 10 international standard strains of Apg (including C-3 strains) were preserved in the laboratory; four domestic isolates and vac-

Table 1. Primers used for this experiment

| Primer	Sequence (5'→3')	Length/bp	Characteristic
Fdp	ACTAATTTATTAGATTCTGATGATATTCGTCAYATGYTNA	996	Deputy chickens against bacillus aroA mergers primers
Rdp	GGTTGAATACGAATAAAATCTTGACCYTCYTCNAC		
FaroA	TTGGGGACACGATTTAGACGGC	1963	Contains aroA sequence primers
RaroA	AATTCCTCAAAGAATGTTGGGAAGG		
Af	GCGATGCTTCTTCTGCCTCTTA	324	Amplification aroA downstream section
Af2	GCGAGGATTTTATCCAAGCAGA		
Af3	GACCGCCTGAGTGCAATGCAA		
Ar	GACGTAATGCGTCCACCAAATG	910	Amplification aroA upstream fragments
Ar2	CTCAAAGGTCGCATCGCCGTT		
Ar3	TCGCATTCACGCTTATCTTGTG		
Fa	ATGGAAAAATTAACGCTACAAC	1293	AroA Gene primer
Ra	TTATCGTACAATCTTCGCAAATT		
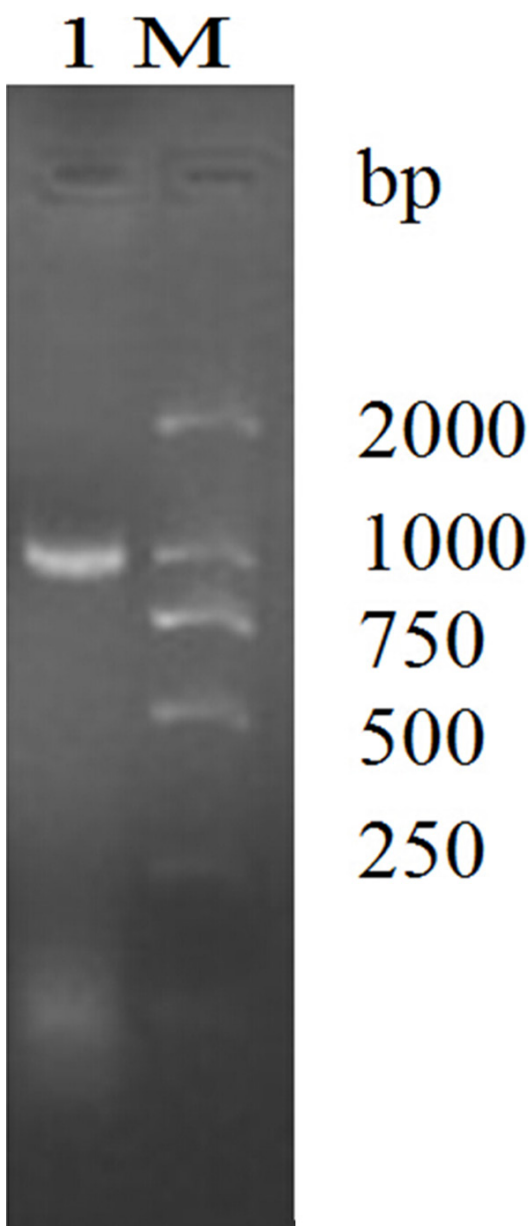
Figure 1. AroA gene partial fragment M. D2000 marker. 1. strain 145.

cine strains were donated by Chen Xiaoling teacher.

Reagents and instruments: TSA and TSB were purchased from BD Biotech Co.LTD; genomic DNA extraction kit was purchased from OMEGA company; PCR mix and D2000 were purchased from Tiangen Biotech company; BSA was purchased from NEB company; gradient PCR instrument (Germany Eppendorf); agarose gel imaging Systems (BioLab Inc.); gene sequencing and primer synthesis were performed by Beijing Liuhe Huada gene technology company.

Software for primers and sequence analysis: All designed primer sequences were shown in **Table 1**. The sequence analysis was carried out by BLAST, DNA STAR, MEGA and EXPASY.

Methods

Genome extraction: The single colony grown on TSA was picked to be cultured in TSB containing calf serum and NAD at 37°C overnight, and 3 ml broth was centrifuged. Genome extraction was performed according to the method provided in the kit.

Amplification of partial: AroA gene Firstly homology analysis of 16sRNA from Apg was performed, and then the amino acid sequences of aroA with high similarity experienced block analysis. Degenerate primers for the aroA gene were designed by CODEHOP. The degenerate primers were used to amplify partial sequences
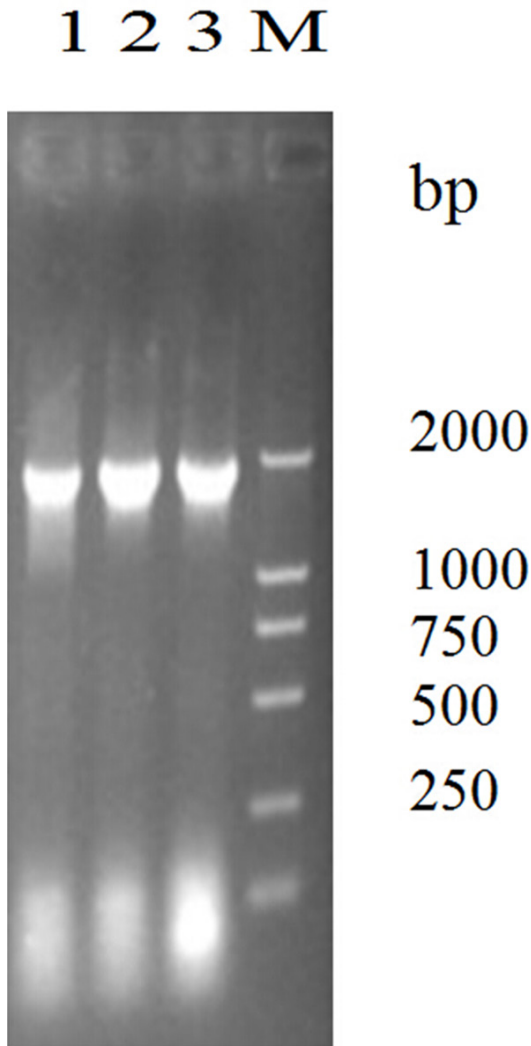
Figure 2. AroA complete sequence 1. Strain 0083. 2. Strain 0222. 3. Strain modest. M. D2000.
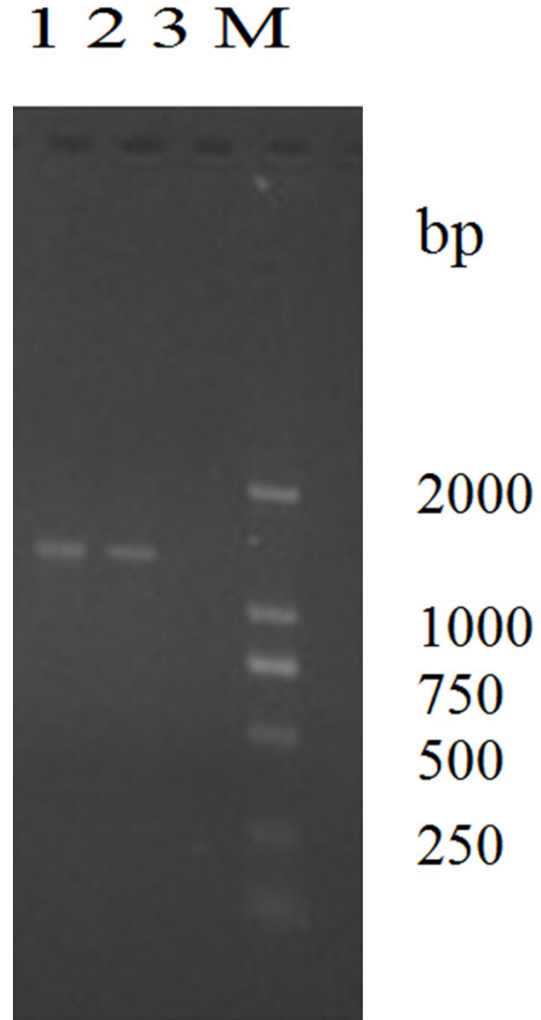

Figure 3. AroA gene verification results. 1. AroA complete sequence. 2. Strain genome. 3. Negative. Control M. D2000 marker.

using the method of touchdown PCR amplification.

TAIL-PCR amplification and sequence assembly: Improved TAIL-PCR method [7] was used to amplify partial aroA gene from upstream to downstream, in order to get the full aroA sequence. AroA gene fragment of 700 bp was used to design two specific nested primers, which were used to amplify its upstream and downstream sequence. The degenerate primers in the walking kit were replaced with 10 bases of RAPD primers; finally specific bands were obtained by thermal asymmetric three nested PCR. The obained fragment was recovered and sequenced to link with the aroA fragment by DNASTAR.

Acquisition and identification of aroA gene: According to the spliced full-length sequence, the upstream and downstream primers of FaroA and RaroA were designed. Three standard strains were amplified to get whole aroA gene. The recovered fragment and genome were taken as a template to reverse identify aroA gene with Ra and Fa as primers. Finally, the spliced gene was confirmed to be aroA gene of Apg by BLAST comparison.

Functional sites of aroA gene and its encoded protein analysis

Functional sites of aroA gene were analyzed and identified by ClustalW2.0, a kind of sequence analysis online tool. And its adjacent
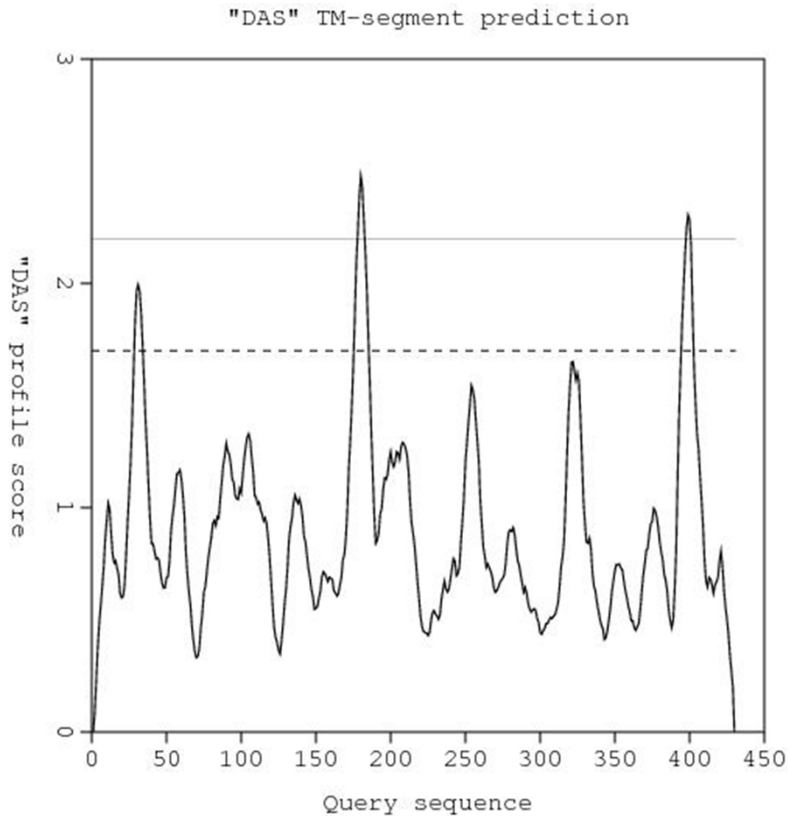
Figure 4. Transmembrane area forecasting results of aroA conding.

Figure 5. AroA amplifying of different serotype 1. Strain 0083. 2. Strain 0222. 3. Strain modest. 5. Strain 668. M. D2000 marker.

sequences between upstream and downstream were compared using BLSAT and sequenced. Finally, the related method published in EXPASY was used to predict the secondary structure of aroA-encoding protein and B cell epitopes.

Amplification and sequence alignment of each serotype of aroA gene

In this study, international standard strains 0083 (A-type), 0222 (B type), modest (C-type) as well as domestic isolates B2 (B type) and vaccine strain 668 (C-type) were selected. DNA was amplified using Fa and Ra as primers of aroA full-length sequences, followed by amino acid sequence analysis and the comparison between encoded products in the various strains.
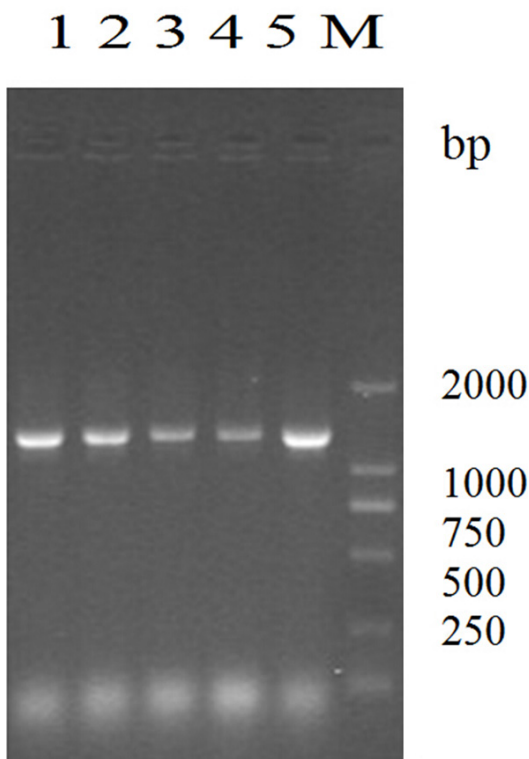
Sequence alignment of aroA gene in the various strains

Whole aroA gene sequences of Haemophilus and Pasteurella were collected from GeneBank. They were analyzed with MEGA4.0; genetic distance matrix and phylogenetic tree were constructed, so as to get the relationships of aroA genetic and evolution among strains.

Results

Acquisition and identification of whole aroA genesequences

Using Fdp and Rdp as degenerate primers, international standard strain 145 was amplified to obtain a partial aroA gene fragment of 996 bp (**Figure 1**); after extending and splicing, the complete aroA ORF sequence of 2178 bp was obtained. Furthermore, according to the splice sequences, the upstream and downstream primers of FaroA and RaroA were designed, and fragments of 1963 bp containing whole aroA gene were obtained by PCR amplification of three standard strains (**Figure 2**). Respectively taking the recovered genom and fragment as templates, Fa and Ra as prim-

Figure 6. The homology of the amino acid sequence of Apg aroA.

ers, the whole aroA gene was verified, and t the amplified specific fragments were consistent with the expected results (**Figure 3**). This gene sequence contained the whole aroA gene fragment (1293 bp). The aroA gene sequence obtained a gene sequence number of JF834213 in GenBank. The protein encoded by the gene was 5'-pyruvate shikimate-3-phosphate synthase (EPSPS); the encoded amino acids had a length of 430aa and a molecular weight of 46.9 KD. BLAST alignments confirmed that the spliced gene was indeed aroA gene of Apg.

AroA gene and its encoding protein analysis

Analysis of functional sites showed that its main functional sites located at: amino acids of 90-104 and 341-359. Upstream fragment of the gene was histidinol phosphate aminotransferase (HisC) gene, with a pI of 5.31; the half-life was greater than 10 h and instability index was 30.18, more stable. N-terminal sequence was Methionine, with 52 negatively charged residues (Asp + Glu) and 44 positively charged residues (Arg + Lys). Instability index was 30.18. EXPASY methods suggested that secondary structure of the protein dominated by flexible regions, and the potential B-cell epitopes were: 49-64, 99-114, 407-422, 322-337, 152-167. DAS results indicated that the protein had

three transmembrane domains (**Figure 4**). According to Burkhard [8] theory, the protein was a hybrid protein. Combining the secondary prediction, the protein was a globular protein with a relatively scattered structure.

Amplification of different serotypes of aroA gene

Taking Fa and Ra as upstream and downstream primers, fragments of the expected size were obtained through PCR amplification of five serotypes (**Figure 5**).

Gene sequence

Genetic distances of aroA encoding products among different serotypes of Apg were shown in **Figure 6**; nucleic acid alignments between different bacteria and phylogenetic trees were respectively shown in **Figures 7** and **8**.

Sequence analysis showed that aroA amino acid sequences of different serotypes of Apg showed a high degree of similarity, which once again verified the conservation of the gene. Nucleic acid alignments between different bacteria showed that excluding Shigella, E.coli, and Bacillus, Apg had a similarity of more than 65% with Salmonella and Yersinia, and the similarity with Pasteurella was more than 75%. In addition, it had a high similarity with the aroA gene of parasuis reported by CHEN Huan [9] and Xue Xiaojing [10] in domestic and the coverage rate was large, suggesting that the genetic evolution of aroA gene had a certain regional characteristic, while phylogenetic tree showed it was a separate one.

Conclusion

This trial has combined the degenerate PCR and improved chromosome walking for the research of genes complete sequence for the first time (conventional chromosome walking techniques have been improved), and obtained the expected results. Meanwhile, bioinformatics methods were applied in the study to analy-

Percent Identity

Divergence		1	2	3	4	5	6	7	8	9	10	11	12		
	1		36.5	67.4	37.4	38.3	36.9	37.4	36.9	36.0	37.2	67.1	35.3	1	Actinobacillus pleuropneumoniae serovar
	2	147.0		39.1	74.3	74.3	74.5	74.8	74.8	74.2	65.4	38.9	65.5	2	Avibacterium paragallinarum strain C-3
	3	43.3	127.8		37.7	39.1	38.9	38.7	38.7	37.7	37.8	95.6	36.4	3	Escherichia coli IAI39 chromosome.meg
	4	143.1	31.5	138.0		77.4	74.2	74.3	75.8	75.3	64.7	37.3	64.4	4	Haemophilus influenzae 10810.meg
	5	132.5	31.6	128.6	27.1		74.6	75.0	75.0	75.0	65.7	38.9	65.3	5	Haemophilus parainfluenzae T3T1.meg
	6	147.9	31.5	131.5	32.1	31.5		98.1	78.3	76.9	66.7	38.9	65.7	6	Haemophilus parasuis 3-phosphoshikimat
	7	143.8	30.9	134.0	32.1	30.1	1.9		78.7	77.3	66.2	38.6	65.8	7	Haemophilus parasuis SH0165.meg
	8	148.3	31.0	132.6	29.5	30.6	25.8	25.2		97.9	65.4	38.5	65.0	8	Mannheimia haemolytica strain PH296.meg
	9	155.2	31.8	139.2	30.3	30.6	27.8	27.2	2.1		64.2	37.3	64.7	9	Pasteurella haemolytica serotype 2.meg
	10	140.6	46.8	136.4	48.3	46.4	44.4	45.3	47.0	49.4		37.2	71.3	10	Salmonella enterica subsp. enterica ser
	11	43.7	129.5	4.5	141.4	129.8	131.9	134.3	134.4	142.6	141.2		36.5	11	Shigella dysenteriae Sd197.meg
	12	154.1	46.8	144.0	49.2	47.2	46.4	46.0	47.8	48.3	36.6	143.2		12	Yersinia pestis A1122.meg
		1	2	3	4	5	6	7	8	9	10	11	12		

Figure 7. The homology of aroA gene nucleotide sequence of the diverse members of Pasteurellaceae and other related bacteria.
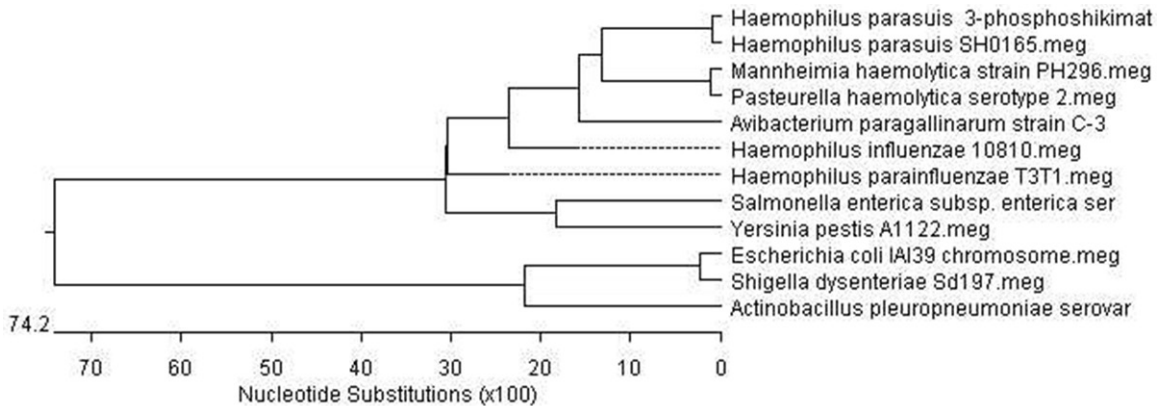

Figure 8. Phylogenetic analysis on the *aroA* gene of Gram-negative organism.

sis and predict the secondary structure of Apg related proteins, Potential B cell epitopes and transmembrane regions. As a result, five potential epitopes B and three transmembrane domains were obtained. Finally, by analyzing the sequence alignment of avibacterium coli serotypes and the same family of bacteria, it further validates the conservative of aroA gene.

Discussion

In order to verify the reproducibility of this method, we cloned another conservative immunogenic protein: membrane protein D-15 using the same method and achieved a success. So this method can provide new ideas and methods for the clone of a variety of virulence associated genes. The hydrophilic parameters, flexibility parameters, accessibility parameters as well as epitope parameters were comprehensively analyzed and compared. It showed that

prediction results obtained by this method had an important reference value to bacterial biology research and vaccine development [11].

Of many Gram-negative bacteria, the upstream of gene aroA is gene serC or phosphoenolpyruvate carboxylase gene [12, 13]. However, in this test, the upstream of aroA gene is part of histidinol phosphate aminotransferase (HisC) gene sequence. This is equal to the latest report that the aroA gene and HisC gene are in the same structure of transcriptional operon [14, 15].

The sequence determination result show that the homology of the amino acid sequence in different serotypes of Apg, which is encoded by aroA gene, is more than 91.4%, indicating a high conservative. And among other members of the aroA gene in Pasteurella Branch, it also has a high homology. This is consistent with the

view of reported literature that aroA gene is more conservative [16]. It provides a theoretical basis for the research of biological functions encoded protein, diagnostic reagents, subunit vaccines and the development of gene deletion mutation.

Acknowledgements

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Hong-Jun Wang, Institute of Animal Husbandry and Veterinary Medicine, Beijing Academy of Agriculture and Forestry Sciences, No. 9, Shuguang Garden Middle Road, Haidian District, Beijing 100097, P. R. China. Tel: +86+010-51503360; E-mail: wanghj_vip@126.com

References

[1] Blackall PJ, Christensen H, Beckenham T, Blackall LL and Bisgaard M. Reclassification of Pasteurella gallinarum, [Haemophilus] paragallinarum, Pasteurella avium and Pasteurella volantium as Avibacterium gallinarum gen. nov., comb. nov., Avibacterium paragallinarum comb. nov., Avibacterium avium comb. nov. and Avibacterium volantium comb. nov. Int J Systemat Evolut Microbiol 2005; 55: 353-362.

[2] Blackall PJ, Yamaguchi T and Matsumoto M. Infectious coryza. In: Calnek BW, Barnes HJ, Beard CW, McDougald LR, Saif YM, editors. Diseases of Poultry. Ames: Iowa State University Press; 1997. pp. 179-190.

[3] Kume K, Sawata A, Nakai T and Matsumoto M. Serological classification of Haemophilus paragallinarum with a hemagglutinin system. J Clin Microbiol 1983; 17: 958-964.

[4] Zhang PJ, Miao M, Sun H, Gong Y and Blackall PJ. Infectious coryza due to Haemophilus paragallinarum serovar B in China. Aust Vet J 2003; 2: 96-97.

[5] Robert R Bragg, Marisa Barbakov, Gutter Gutter and Avner Finger. High virulence vice chickens coli serotype C-3 strains spreading in South Africa. Proceedings of the 15th World Conference on Avian Disease 2007.

[6] Song MX, Chen XL, Mu JQ, Terry T, Wang LL, Li F, Wai W and Zhang XM. AroA gene is used to establish the chicken infectious rhinitis PCR diagnostic method. Zhong Guo Shou Yi Ke Ji 2001; 31: 7-9.

[7] Qiu YG, Tian JH, Ge RC, Zhao BC, Shen YZ and Huang ZJ. A modified TAIL-PCR and its application in isolating gene promoter of wheat. Sheng Wu Gong Cheng Xue Bao 2008; 24: 695-699.

[8] Burkhard R. Predicting one-dimensional protein structure by profile based neural networks. Methods Enzymol 1996; 266: 525-539.

[9] Xu Z, Yue M, Zhou R, Jin Q, Fan Y, Bei W and Chen H. Genomic characterization of Haemophilus parasuis SH0165, a highly virulent strain of serovar 5 prevalent in China. PLoS One 2011; 6: E19631.

[10] Xue X, Xu F, Shi A, Zhang P, Chen X, Yang B and Wang J. Identification and phylogenetic analysis on the aroA gene of Haemophilus parasuis. Wei Sheng Wu Xue Bao 2008; 48: 1100-3.

[11] Sette A and Fikes J. Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. Curr Opin Immunol 2003; 15: 461-70.

[12] Duncan K, Coggins R. The serC-aroA operon of Escherichia coli. Biochem 1986; 234: 49-57.

[13] Garside LH, Collins M, Langford PR, Rycroft AN. Actinobacillus pleuropneumomiae serotype 1 carrying the defined aroA mutation is fully avirlent in the pig. Res Veter Sci 2002; 72: 163-167.

[14] Hong SH, Kim JS, Lee SY, In YH, Choi SS, Rih JK, Kim CH, Jeong H, Hur CG and Kim JJ. The genome sequence of the capnophilic rumen bacterium Mannh-eimia Succiniciproducens. Nat Biotechnol 2004; 22: 1275-1281.

[15] Foote SJ, Bossé JT, Bouevitch AB, Langford PR, Young NM and Nash JH. The complete genome sequence of Actinobacillus pleuropneum-oni-ae L20 (serotype 5b). J Bacteriol 2008; 190: 1495-1496.

[16] Cascon SA, Anguita CJ, Hernanz MC, Sanchez SM, Yugueros J, Naharro G. RFLP-PCR analysis of the aroA gene as a taxonomic tool for the genus Aeromonas. FEMS Microbiol Lett 1997; 156: 199-204.