

ajava

Asian Journal of Animal and Veterinary Advances



Academic
Journals Inc.

www.academicjournals.com

Genetic Phylogenetic Analysis of PRRSV ORF5 Gene Based on a Geometrical Tool

Hua Tang, Xin-Sheng Liu, Yu-Zhen Fang, Li Pan, Zhong-Wang Zhang, Peng Zhou, Jian-Liang Lv, Shou-Tian Jiang, Wen-Fa Hu, Pan Zhang, Yong-Guang Zhang and Yong-Lu Wang

State Key Laboratory of Veterinary Etiological Biology, National Foot and Mouth Disease Reference Laboratory, Key Laboratory of Animal Virology of Ministry of Agriculture, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, China

*Corresponding Author: Yong-Lu Wang and Yong-Guang Zhang, State Key Laboratory of Veterinary Etiological Biology, National Foot and Mouth Disease Reference Laboratory, Key Laboratory of Animal Virology of Ministry of Agriculture, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, China
Tel/Fax: +86-931-8343796*

ABSTRACT

The Z-curve is a geometrical tool for visualizing and comparing genomes. Since, the curve contains the information carried by the given sequence, DNA sequences could be analyzed systematically. In this study, the ORF5 gene of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) was analyzed by the Z-curve method and the DNASTar (DNASTAR Inc.) computer programs. The results revealed by these two methods were identical. The phylogenetic tree is more digitized and the Z-curve method is more picturesque. It turned out to be that, the Z-curve method shows a widely application prospect in phylogenetic relationships analysis. However, the Z-curve method is still in its premature stage. Novel algorithms are expected to be exploited to extract more information involved in the Z-curves.

Key words: Z-curve, geometrical study, PRRSV, ORF5, genetic diversity

INTRODUCTION

Analyzing DNA sequences and amino acid sequences with the aid of geometric approach is a novel way which shows the merits of intuitive, visual, vivid (Chu *et al.*, 2009). Some people have applied geometric approaches into study, such as Li *et al.* (2010) and Giacomo *et al.* (2007).

According to Zhang and Zhang (2003) and Wen and Zhang (2003), Z-curve is a geometric tool to study the DNA sequence. Since, it was first proposed, the Z-curve has been extensively applied to life science research, such as, sequence segmentation (Zhang and Zhang, 2003; Wen and Zhang, 2003), horizontal gene transfer detection (Zhang and Zhang, 2003), isochoric domain inference (Zhang and Zhang, 2003; Wen and Zhang, 2003) and sequence analysis, for instance, nucleotide distribution analysis (Ou *et al.*, 2003), replication origins Identification (Zhang and Zhang, 2005), protein coding genes recognition (Zhang and Wang, 2000), separate base usages of genes revelation (Guo and Yu, 2007) and so on.

Porcine Reproductive and Respiratory Syndrome (PRRS) disease is caused by a positive single-stranded polyadenylated RNA virus (PRRSV) which belongs to Arteriviridae virus family in the order Nidovirales (Conzelmann *et al.*, 1993). And its genome contains eight Open Reading

Frames (ORFs) which is about 15 kD (Cavanagh, 1997). Bred gilts and sows infected PRRSV may display abortion and premature farrowing and the nosogenic liveborn piglets may show respiratory tract disease (Cheon and Chae, 1999).

The PRRSV ORFs were consisted of 1a, 1b, 2B7 (Grebennikova *et al.*, 2004). ORF1a and 1b are believed to encode nonstructural proteins, involved with viral replication and transcription (Grebennikova *et al.*, 2004). The ORFs 2B7 encode the PRRSV structural proteins (Grebennikova *et al.*, 2004). An effective way to distinguish wild type and vaccine-like strains of PRRSV is analyzing the Restriction Fragment Length Polymorphism (RFLP) of ORF 5 (Gagnon and Dea, 1998; Wesley *et al.*, 1998; Cheon and Chae, 2000). Restriction sites diversities are awfully useful in diagnosis of disease (Cheon and Chae, 2004). GP5, the major viral envelope protein, encoded by ORF 5 is the most variable viral structural protein (Mardassi *et al.*, 1995) and it is relevant to virus neutralization (Zhou *et al.*, 2009). A variety of vaccines associating with GP5 were developed and performed well (Duran *et al.*, 1997; Qiu *et al.*, 2005; Xue *et al.*, 2004).

Gaining insight into the information carried by the genomes of PRRSV would be helpful to discover drugs, develop vaccines as well as monitor their evolution and spread (Grebennikova *et al.*, 2004). According to the genetic analysis of PRRSV isolates, PRRSV was divided into two predominant genotypes, the European and North American genotypes (Grebennikova *et al.*, 2004; Lee *et al.*, 2006). The ORF5 gene of PRRSV have been widely analyzed and researchers found it was useful to help us understand the genetic diversity of PRRSV isolates (Cha *et al.*, 2006; Chen *et al.*, 2006; Indik *et al.*, 2000, 2005; Mateu *et al.*, 2006; Stadejek *et al.*, 2006; Thanawongnuwech *et al.*, 2004; Zhou *et al.*, 2009). In this study, the complete sequences of the ORF5 gene of the twenty-one PRRSV isolates were analyzed with the DNASTar (DNASTAR Inc.) computer programs and the Z-curve method.

MATERIALS AND METHODS

The PRRSV ORF5 genes and annotated information (Table1) were downloaded at the web site (<http://www.ncbi.nlm.nih.gov/>).

The Z-curve coordinates compute software (Zplotter) was downloaded at the web site (<http://tubic.tju.edu.cn/zcurve/>). Users are recommended to use Z curve plotter Java applet version which can be available at the web site (<http://tubic.tju.edu.cn/zcurve/>). However, for those who experience difficulty in using Java applet, or want to compute Z-curve coordinates locally, this local version of Zplotter can be used. The theory of Z-curve was reported by Zhang and Zhang (2003).

RESULTS

Twenty-one PRRSV isolates ORF5 gene were performed genetic phylogenetic analysis to reveal the genetic relationship and evolution of these PRRSV isolates with the help of DNASTar program (Fig. 1). Six distinct clusters were showed in the phylogenetic tree and we marked them A, B, C, D, E, F, respectively. And it is obviously that, there are two apparent genotypes among these twenty-one PRRSV isolates.

The twenty-one PRRSV isolates previously studied were then analyzed by the Z-curve method (Fig. 2). Figure 2a shows two evident genotypes of these virus the same as the result displayed in the phylogenetic tree. Figure 2b-d, respectively mapped the A cluster; B cluster and C, D, E, F

Table 1: The Z-curve plot marks, accession numbers, isolation regions of collected PRRSV and virus isolates cluster categories distinguished in the phylogenetic tree

Z-curve plot mark	Accession number	Isolation region	Cluster
'e:\5\1.txt' u 1:2:3	AB546121.1	Japan	A
'e:\5\2.txt' u 1:2:3	AF095503.1	Denmark	A
'e:\5\3.txt' u 1:2:3	AF095518.1	Denmark	A
'e:\5\4.txt' u 1:2:3	AF176426.1	USA	A
'e:\5\5.txt' u 1:2:3	AF176427.1	USA	A
'e:\5\6.txt' u 1:2:3	AF339493.1	USA	B
'e:\5\7.txt' u 1:2:3	AF339494.1	USA	B
'e:\5\8.txt' u 1:2:3	AF339495.1	USA	B
'e:\5\9.txt' u 1:2:3	AF339496.1	USA	B
'e:\5\10.txt' u 1:2:3	AY740010.1	Italy	C
'e:\5\11.txt' u 1:2:3	EU273703.1	Taiwan	D
'e:\5\12.txt' u 1:2:3	JF730974.1	Spain	C
'e:\5\13.txt' u 1:2:3	JN651728.1	Germany	C
'e:\5\14.txt' u 1:2:3	JN651741.1	Germany	B
'e:\5\15.txt' u 1:2:3	JN651746.1	Germany	A
'e:\5\16.txt' u 1:2:3	JQ917911.1	China	F
'e:\5\1.txt' u 1:2:3	U66379.1	USA	B
'e:\5\18.txt' u 1:2:3	U66381.1	USA	B
'e:\5\19.txt' u 1:2:3	U66387.1	Canada	A
'e:\5\20.txt' u 1:2:3	U66392.1	USA	E
'e:\5\21.txt' u 1:2:3	U66399.1	USA	A

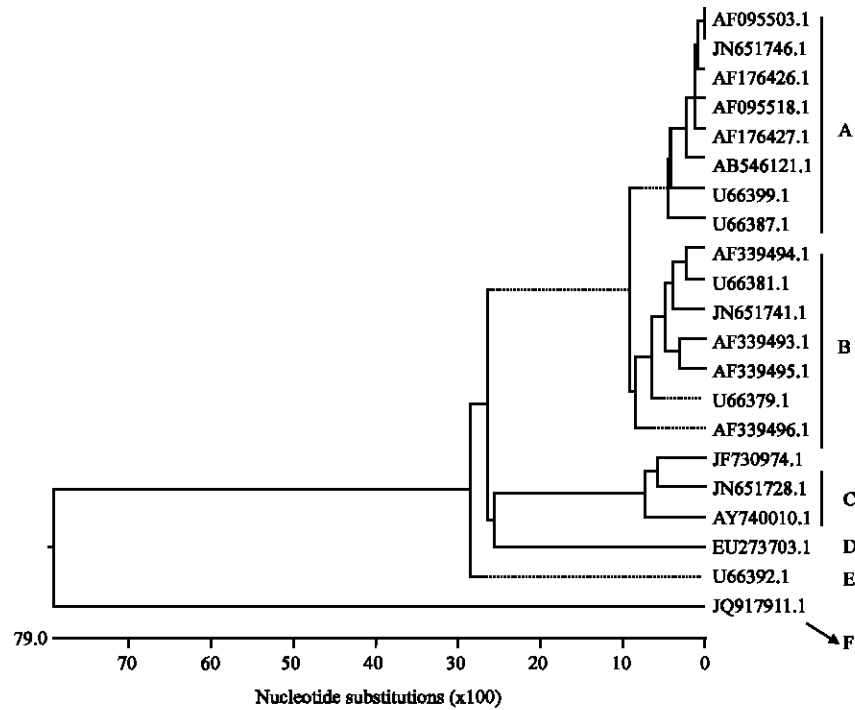


Fig. 1: A phylogenetic tree based on the nucleotide sequence of the ORF5 gene of 21 PRRSV isolates, The tree was constructed with the aid of the DNASTAR program, There are two apparent genotypes and six distinct clusters

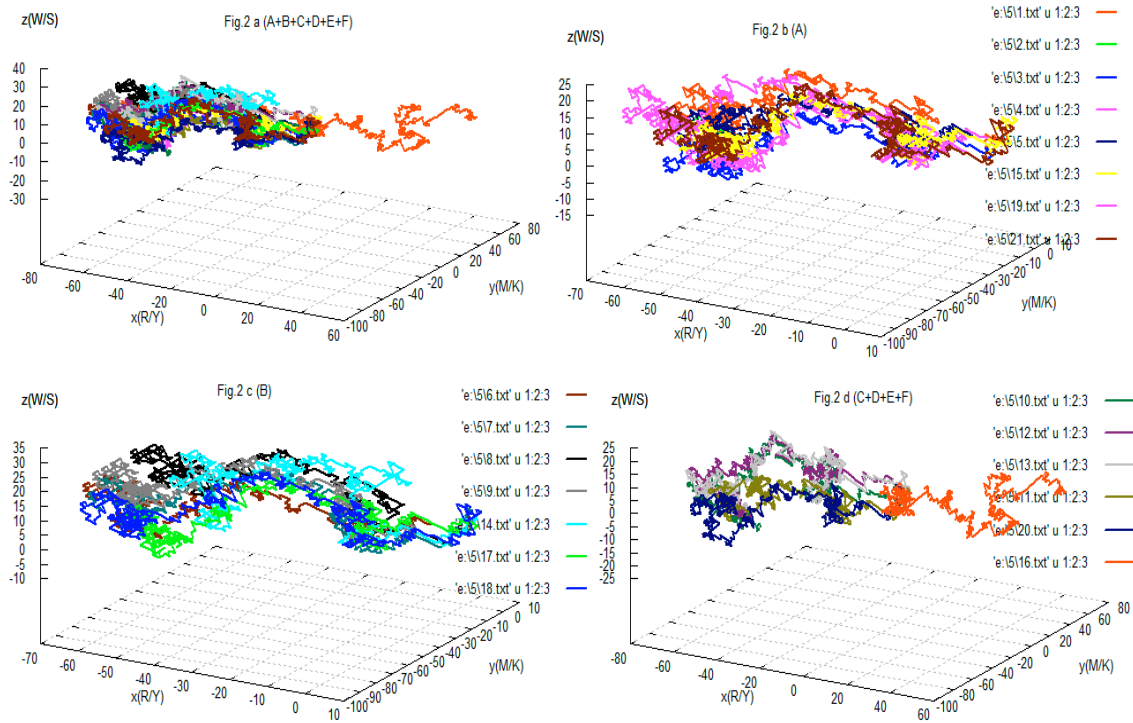


Fig. 2(a-d): The relevant plot marks for PRRSV isolates ORF5 gene used in this graph are listed in Table 1 (a) Portrays the three-dimensional Z-curves of all twenty-one PRRSV isolates ORF5 gene (AB546121.1, AF095503.1, AF095518.1, AF176426.1, AF176427.1, AF339493.1, AF339494.1, AF339495.1, AF339496.1, AY740010.1, EU273703.1, JF730974.1, JN651728.1, JN651741.1, JN651746.1, JQ917911.1, U66379.1, U66381.1, U66387.1, U66392.1, U66399.1), (b-d), respectively portray the ORF5 gene of PRRSV isolates clusters (A; B; C, D, E, F) distinguished in the phylogenetic tree (Fig. 1), (b) Z-curves of AB546121.1, AF095503.1, AF095518.1, AF176426.1, AF176427.1, JN651746.1, U66387.1, U66399.1, (c) Z-curves of AF339493.1, AF339494.1, AF339495.1, AF339496.1, JN651741.1, U66379.1, U66381.1 (d) Z-curves of AY740010.1, JF730974.1, JN651728.1, EU273703.1, U66392.1, JQ917911.1. That is to say, all the curves plotted in Fig. 2a are the union of curves graphed in Fig. 2-d. The legends of Fig. 2a are the union of legends plotted in Fig. 2b-d. As can be observed in Fig. 2a, the conspicuous red curve is the Z-curve of isolate JQ917911.1. It is obviously that, there are two apparent genotypes. From Fig. 2b-d, We can see that the differences between each Z-curve are remarkable. There is no doubt that the results revealed by Z-curve were identical with the consequences disclosed by the DNASTAR program

cluster PRRSV isolates which were distinguished in the phylogenetic tree. There is no doubt that the results revealed by these two methods were identical. The phylogenetic tree is more digitized and the Z-curve method is more picturesque.

DISCUSSION

We can study the molecular phylogenies which may reveal many aspects of the transmission, epidemiology and evolution of rapidly evolving pathogens through genome analyzing (Pybus and Rambaut, 2009). Viral diversity, modulated by natural selection, global patterns of virus circulation and host population biology, is generated by evolutionary dynamics which resulted from a complex combination of rapid mutation and genome segment reassortment (Baillie *et al.*, 2012). So the gene sequence study plays a key role in the life science.

It turned out to be that PRRSV could be gene typed through analyzing the ORF5 gene which was in agreement with Greiser-Wilke *et al.* (2010). Two genotypes of PRRSV were previously reported: American- and European-type strains (Indik *et al.*, 2005; Nelsen *et al.*, 1999; Zhou *et al.*, 2011). And this was in accordance with the results of this study. And it was obviously that the ORF5 gene of PRRSV is highly variable (An *et al.*, 2011). The DNASTAR program and the Z-curve method in this paper showed a identical accuracy in the genetic phylogenetic analysis. The former is more digitized and the Z-curve method is more pictorial. However, the phylogenetic tree could discriminate the genetic diversity more easily which can be proved in the pictures. If we want to confirm a virus belongs to which genotypes according to the Z-curve method we should accomplish a preparation work. The precondition is that we should construct a data base which definitely records the Z-curve coordinates area of different genotypes. Thus, if we plot a virus genome in a clearly marked genotypes Z-curve rectangular space coordinate systems, we could immediately read out what kind of genotypes the targeted virus is. So the Z-curve method is another reliably phylogenetic analysis tool. In addition, the Z-curve also can be used to recognize protein coding genes. Z-curve-E, a new self-training system derived from Z-curve, has been used to recognize protein coding genes (Guo and Zhang, 2006).

In general, Z-curve method shows its widely application in the areas of life sciences. For instance, gene sequence analysis, molecular epidemiology analysis, homology analysis, evolution analysis, phylogenetic relationships analysis, genetic diversity analysis. Currently, the Z-curve method is still in its premature stage. Since the more information is extracted from the Z-curve, the more accurate result can be gained, we are expected to explore more about the Z-curve. Gaining insight into the information carried by the genomes of viruses would be helpful to discover drugs, develop vaccines as well as monitor their evolution and spread. Some other novel algorithms are expected to be exploited to extract more information contained in the Z-curves.

ACKNOWLEDGMENTS

This study was financially supported by the National Pig Industrial System (CARS-36-06B) and Ministry of Agriculture' Special Funds for Scientific Research on Public Causes, Research and Demonstration on evaluation technologies of clinical immune responses of serious animal diseases vaccine, No. 201203039).

REFERENCES

- An, T.Q., Z.J. Tian, Y.J. Zhou, Y. Xiao and J.M. Peng *et al.*, 2011. Comparative genomic analysis of five pairs of virulent parental/attenuated vaccine strains of PRRSV. *Vet. Microbiol.*, 149: 104-112.
- Baillie, G.J., M. Galiano, P.M. Agapow, R. Myers and R. Chiam *et al.*, 2012. Evolutionary dynamics of local pandemic H1N1/09 influenza lineages revealed by whole genome analysis. *J. Virol.*, 86: 11-18.

- Cavanagh, D., 1997. Nidovirales: A new order comprising coronaviridae and arteriviridae. *Arch. Virol.*, 142: 629-633.
- Cha, S.H., E.J. Choi, J.H. Park, S.R. Yoon and J.Y. Song *et al.*, 2006. Molecular characterization of recent Korean Porcine Reproductive and Respiratory Syndrome (PRRS) viruses and comparison to other Asian PRRS viruses. *Vet. Microbiol.*, 117: 248-257.
- Cheon, D.S. and C. Chae, 1999. Distribution of a Korean strain of porcine reproductive and respiratory syndrome virus in experimentally infected pigs, as demonstrated immunohistochemically and by *in-situ* hybridization. *J. Comp. Pathol.*, 120: 79-88.
- Cheon, D.S. and C. Chae, 2000. Restriction fragment length polymorphism analysis of open reading frame 5 gene of porcine reproductive and respiratory syndrome virus isolates in Korea. *Arch. Virol.*, 145: 1481-1488.
- Cheon, D.S. and C. Chae, 2004. Comparison of the pathogenicity of two strains (wild type and vaccine-like) of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in experimentally infected sows. *J. Comp. Pathol.*, 130: 105-111.
- Chen, J., T. Liu, C.G. Zhu, Y.F. Jin and Y.Z. Zhang, 2006. Genetic variation of Chinese PRRSV strains based on ORF5 sequence. *Biochem. Genet.*, 44: 425-435.
- Chu, J.H., S.T. Weiss, V.J. Carey and B.A. Raby, 2009. A graphical model approach for inferring large-scale networks integrating gene expression and genetic polymorphism. *BMC. Syst. Biol.* Vol. 33.
- Conzelmann, K.K., N. Visser, P. van Woensel and H.J. Thiel, 1993. Molecular characterization of porcine reproductive and respiratory syndrome virus, a member of the arterivirus group. *Virology*, 193: 329-339.
- Duran, P.J., I. Climent, J. Sarraseca, A. Urniza, E. Cortes, C. Vela and J.I. Casal, 1997. Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain Olot/91. Involvement of ORF3 and ORF5 proteins in protection. *Virus Genes*, 14: 19-29.
- Gagnon, C.A. and S. Dea, 1998. Differentiation between porcine reproductive and respiratory syndrome virus isolates by restriction fragment length polymorphism of their ORFs 6 and 7 genes. *Can. J. Vet. Res.*, 62: 110-116.
- Giacomo, F., M.M. Francesco, C. Davide and H. Muller, 2007. Graph-based identification of cancer signaling pathways from published gene expression signatures using PubLiME. *Nucleic Acids Res.*, 35: 2343-2355.
- Grebennikova, T.V., D.F. Clouser, A.C. Vorwald, M.I. Musienko and W.L. Mengeling *et al.*, 2004. Genomic characterization of virulent, attenuated and revertant passages of a North American porcine reproductive and respiratory syndrome virus strain. *Virology*, 321: 383-390.
- Greiser-Wilke, I., K. Fiebig, C. Drexler and E.G. Beilage, 2010. Genetic diversity of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in selected herds in a pig-dense region of North-Western Germany. *Vet. Microbiol.*, 143: 213-223.
- Guo, F.B. and C.T. Zhang, 2006. ZCURVE_V: A new self-training system for recognizing protein-coding genes in viral and phage genomes. *BMC Bioinformatics*, Vol. 7.
- Guo, F.B. and X.J. Yu, 2007. Separate base usages of genes located on the leading and lagging strands in *Chlamydia muridarum* revealed by the Z curve method. *BMC Genomics*, Vol. 8.
- Indik, S., L. Valicek, D. Klein and J. Klanova, 2000. Variations in the major envelope glycoprotein GP5 of Czech strains of porcine reproductive and respiratory syndrome virus. *J. Gen. Virol.*, 81: 2497-2502.

- Indik, S., F. Schmoll, W. Sipos and D. Klein, 2005. Genetic variability of PRRS virus in Austria: Consequences for molecular diagnostics and viral quantification. *Vet. Microbiol.*, 107: 171-178.
- Lee, C., D. Hodgins, J.G. Calvert, S.K.W. Welch, R. Jolie and D. Yoo, 2006. Mutations within the nuclear localization signal of the porcine reproductive and respiratory syndrome virus nucleocapsid protein attenuate virus replication. *Virology*, 346: 238-250.
- Li, X., C. Panea., C.H. Wiggins, V. Reinke and C. Leslie, 2010. Learning graph-mer motifs that predict gene expression trajectories in development. *PLoS Comput. Biol.* Vol. 6.
- Mardassi, H., S. Mounir and S. Dea, 1995. Molecular analysis of the ORFs 3 to 7 of porcine reproductive and respiratory syndrome virus, Quebec reference strain. *Arch. Virol.*, 140: 1405-1418.
- Mateu, E., I. Diaz, L. Darwich, J. Casal, M. Martin and J. Pujols, 2006. Evolution of ORF5 of Spanish porcine reproductive and respiratory syndrome virus strains from 1991 to 2005. *Virus Res.*, 115: 198-206.
- Nelsen, C.J., M.P. Murtaugh and K.S. Faaberg, 1999. Porcine reproductive and respiratory syndrome virus comparison: Divergent evolution on two continents. *J. Virol.*, 73: 270-280.
- Ou, H.Y., F.B. Guo and C.T. Zhang, 2003. Analysis of nucleotide distribution in the genome of streptomyces coelicolor A3(2) using the Z curve method. *FEEBS. Lett.*, 540: 188-194.
- Pybus, O.G. and A. Rambaut, 2009. Evolutionary analysis of the dynamics of viral infectious disease. *Nat. Rev. Genet.*, 10: 540-550.
- Qiu, H.J., Z.J. Tian, G.Z. Tong, Y.J. Zhou, J.Q. Ni, Y.Z. Luo and X.H. Cai, 2005. Protective immunity induced by a recombinant pseudorabies virus expressing the GP5 of porcine reproductive and respiratory syndrome virus in piglets. *Vet. Immunol. Immunopathol.*, 106: 309-319.
- Stadejek, T., M.B. Oleksiewicz, D. Potapchuk and K. Podgorska, 2006. Porcine reproductive and respiratory syndrome virus strains of exceptional diversity in Eastern Europe support the definition of new genetic subtypes. *J. Gen. Virol.*, 87: 1835-1841.
- Thanawongnuwech, R., B. Thacker, P. Halbur and E.L. Thacker, 2004. Increased production of proinflammatory cytokines following infection with porcine reproductive and respiratory syndrome virus and *Mycoplasma hyopneumoniae*. *Clin. Diagn. Lab. Immunol.*, 11: 901-908.
- Wen, S.Y. and C.T. Zhang, 2003. Identification of isochore boundaries in the human genome using the technique of wavelet multiresolution analysis. *Biochem. Biophys. Res. Commun.*, 311: 215-222.
- Wesley, R.D., W.L. Mengeling, K.M. Lager, D.F. Clouser, J.G. Landgraf and M.L. Frey, 1998. Differentiation of a porcine reproductive and respiratory syndrome virus vaccine strain from North American field strains by restriction fragment length polymorphism analysis of ORF 5. *J. Vet. Diagn. Invest.*, 10: 140-144.
- Xue, Q., Y.G. Zhao, Y.J. Zhou, H.J. Qiu and Y.F. Wang *et al.*, 2004. Immune responses of swine following DNA immunization with plasmids encoding porcine reproductive and respiratory syndrome virus ORFs 5 and 7 and porcine IL-2 and IFN α . *Vet. Immunol. Immunopathol.*, 102: 291-298.
- Zhang, C.T. and J. Wang, 2000. Recognition of protein coding genes in the yeast genome at better than 95% accuracy based on the Z curve. *Nucleic Acids Res.*, 28: 2804-2814.
- Zhang, C.T. and F. Zhang, 2003. An isochore map of the human genome based on the Z curve method. *Gene.*, 317: 127-135.

- Zhang, R. and C.T. Zhang, 2005. Identification of replication origins in archaeal genomes based on the Z-curve method. *Archaea*, 1: 335-346.
- Zhou, Y.J., H. Yu, Z.J. Tian, G.X. Li and X.F. Hao *et al.*, 2009. Genetic diversity of the ORF5 gene of porcine reproductive and respiratory syndrome virus isolates in China from 2006 to 2008. *Virus Res.*, 144: 136-144.
- Zhou, Z., J. Ni, Z. Cao, X. Han and Y. Xia *et al.*, 2011. The epidemic status and genetic diversity of 14 highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV) isolates from China in 2009. *Vet. Microbiol.*, 150: 257-269.