African Journal of Biotechnology Vol. 11(18), pp. 4305-4310, 1 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.3550
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Sequences polymorphism and variation of major histocompatibility complex DRB exon 2 of black Dahe pig

Lizhou Tang[1], Long Yu[1], Jiangang Chen[1], Junjie Wang[1], Mei Ma[1], Weidong Lu[1*] and Tongzuo Zhang[2*]

[1]Yunnan-Guizhou Plateau Institute of Biodiversity, College of Biologic Resource and Environment Science, Qujing Normal University, Qujing, Yunnan 655011, China.
[2]Key Laboratory of the Qinghai-Tibetan Plateau Ecosystem and Biological Evolution and Adaptation, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai 810001, China.

Accepted 17 January, 2012

DNA sequences of swine leukocyte antigens (SLA) DRB exon 2 from five populations were used to investigate genetic polymorphism of black Dahe pig in Yunnan province of China. Our results showed that the SLA-DRB exon 2 of black Dahe pig obtained high variable rate of 23.50%, which was higher than the values analyzed for cattles or other livestocks. The average G + C content at the third synonymous variable coding positions (GC_3) was closed to 0.78. The amino acids composition rate and majority of relative synonymous codon usage were near the ranges of 0.00 to 11.27 and 1.00 to 2.00, respectively. These significant polymorphism, high GC_3 content and differentiation of preferred codon usage could be resulted from over dominance selection, base compositional bias and the diversity of gene expression level. Aforementioned results adequately testified the SLA-DRB exon 2 may be an important genetic marker for studies in genetic diversity, population genetics, breeding reproduction investigation and exploring new functional genes of black Dahe pig.

Key words: Black Dahe pig, DRB exon 2, genetic polymorphism, major histocompatibility complex, swine lymphocyte alloantigen, Yunnan province of China.

INTRODUCTION

Major histocompatibility complex (MHC) gene is a member of the multigene families, which include three gene regions (Classes I, II and III) and correlate with immunity function in vertebrate (Benacerraf, 1981; Gu and Nei, 1999). Recent studies showed that the compact chain, differ markedly in degree of expression and high polymorphism were the characters for some MHC genes, such as exon 2 of MHC-DRB and MHC-DQB (Kloch et al., 2010). Previous researches of MHC gene mostly concentrated their attention on rats, and found that this gene was also discovered in different chromosomes of more vertebrates, such as sheeps and cattles (Gorer, 1937). Some studies indicated MHC gene plays important role in the regulation of immune response, also could be related to the characteristics of disease resistance, reproductive capability, growth capability and fleshy quality in livestock and poultry (Benacerraf, 1981). Therefore, as an important candidate gene for investigation about disease resistance and susceptibility, the MHC gene has become a hotspot in the research area of genetic breeding, and has attracted the attention of some scientists, such as ecologists, geneticists and animal breeding scientists.

The MHC gene of pigs, named swine lymphocyte alloantigen (SLA), have been originally pitched on chromosome number 7 and have been finally located on the centromere of chromosome number 7 (Rabin et al., 1985). Other studies mostly investigated the relationship between SLA genotypes and immunoreactions or disease

*Corresponding author. E-mail: luweidonghxx@yahoo.com.cn, zhangtz@nwipb.ac.cn. Tel: +86-874-8998627. Fax:: +86-874-8998627.

Table 1. The locations of sampled populations of black Dahe pig.

Population	Site	Latitude (N)	Longitude (E)	Sample size
1	Dahe town, Fuyuan county, Yunnan province	25°32'51.79"	104°18'10.16"	10
2	Yingshang town, Fuyuan county, Yunnan province	25°28'37.52"	104°19'33.40"	10
3	Zhongan town, Fuyuan county, Yunnan province	25°41'29.73"	104°14'52.72"	10
4	Fucun town, Fuyuan county, Yunnan province	25°22'23.36"	104°26'29.68"	10
5	Laochang town, Fuyuan county, Yunnan province	25°13'34.50"	104°31'15.63"	8

resistance (Benacerraf, 1981; Shia et al., 1995). More scientists subsequently studied differentiation of parasites resistance of SLA genotypes in different pigs or tissues, and analyzed relationship between some economic characters and gene haplotypes based on molecular biologic techniques (Lee et al., 2005). As a crossbreed hybridized by Chinese Dahe pig and England Yorkshire pig, black Dahe pig has been formally approved as an endemic new species by the National Agriculture Ministry of China, and distributed on districts of Dahe-Yingshang towns in Qujing city of Yunnan province. This species has become the material source for Chinese famous ham brand (Xuanwei ham) because of its excellent characters, such as better meat, high intramuscular fat, quick growth rate, strong reproductive ability, resistance of crude feed and disease resistance (Huo et al., 2009). However, the piglets of black Dahe pig were also killed by these familiar epidemic diseases, especially the yellow and white scour of piglets (Huo et al., 2009). The clinic diseases could badly affect the pure reproduction, piglet supply and piglet production with high quality, and bring huge economic loses to the farmer. The medication therapy or vaccine inoculability could not settle the difficult problem of repetitive disease infection ultimately. But pig producing via the method of gene modification must satisfy modern people's demands for food safety.

Therefore, we analyzed the sequence, amino acid substitution and polymorphism of exon 2 of SLA-DRB gene for endemic black Dahe pig from Yunnan province of China. The aforementioned acquired information can facilitate us to develop subsequent studies of disease resistance and breeding, and also to investigate the relationship between other economic characters and SLA genotypes of this endemic breed.

MATERIALS AND METHODS

Population samples

Samples for genetic polymorphism analysis totaled 48 individuals from 5 populations of black Dahe pig were collected and sequenced for this study (Table 1). The sampled muscle tissues for each pig were immediately preserved in 95% ethanol and transferred to Qujing Normal University, for storage at -20°C.

DNA extraction, PCR amplification and sequencing

Total DNA was isolated from ethanol-fixed muscle tissue after proteinase K digestion, followed by standard animal genome DNA extraction (Dingguo, Beijing, China) and ethanol precipitation. The sequences of exon 2 of MHC-DRB gene were amplified using primer pair DRB1 (5'-TAGGATCCCTCACAGCGCATTTCTT-3') and DRB2 (5'-GTGTCTGCAGTACGTGTCCA-3'). PCR amplifications were performed in total reaction volumes of 15 µl, containing 1.5 µl of 10 × Buffer (include Mg^{2+}), 1 µl of mixed dNTP, 1.2 µl of each primer (synthesized by Sangon, Shanghai, China), 1.5 µl (100 ng/µl) of template DNA, and 0.12 µl *Taq* DNA polymerase (5 U/µl) (Sangon, Shanghai, China). The reaction mixtures were denatured at 94°C for 5 min and subjected to 30 cycles of 30 s at 94°C, 45 s at 60°C, 45 s at 72°C, and a final extension step of 8 min at 72°C. PCR products were purified using a Geneclean Purification Kit, cloned and sequenced (Xu et al., 2005; Wu et al., 2007).

Data analysis

All obtained sequences for black Dahe pig were compared with the submitted sequences of *Sus scrofa* in GenBank (accession numbers GU263819, U52528). All sequences were aligned using Clustal X (Thompson et al., 1997) with the default settings and refined manually. The total number of mutations, sites with alignment gaps, polymorphic sites, singleton variable sites and parsimony informative sites were determined using DnaSP (Rozas et al., 2003). Maximum likelihood estimates of transition/transversion bias were conducted in MEGA5 (Tamura et al., 2011) based on using a user-specified topology and eliminating all missing positions. The nucleotide frequency, G + C content, relative synonymous codon usage, average nucleotide composition, amino acid composition and average nucleotide pair frequency were estimated in MEGA5 (Tamura et al., 2011).

RESULTS

Homologous comparison and clone sequencing

As compared with exon 2 of SLA-DRB DNA sequences of

Table 2. The maximum composite likelihood estimate of the pattern of nucleotide substitution.

Nucleotide	A	T	C	G
A	-	*5.04*	*5.71*	**12.94**
T	*5.48*	-	**15.86**	*7.86*
C	*5.48*	**14**	-	*7.86*
G	**9.01**	*5.04*	*5.71*	-

Rates of different transitional substitutions are shown in **bold** and those of transversional substitutions are shown in *italics*. The nucleotide frequencies are 22.73% (A), 20.93% (T/U), 32.64% (C), and 23.71% (G). The transition/transversion rate ratios are k_1 = 1.645 (purines) and k_2 = 2.778 (pyrimidines). The overall transition/transversion bias is R = 1.052, where R = [A*G*k_1 + T*C*k_2]/[(A+G)*(T+C)].

Sus scrofa, the homologous percentage for all obtained black Dahe pig sequences were close to 90 to 100%. A total of 29 haplotypes were identified in 48 pig individuals. There are 200 sites determined using clone sequencing.

Sequence polymorphism and variation

Of the 200 sites aligned in all individuals, 5 missing sites, 148 invariable sites and 47 polymorphic sites were observed respectively. These 47 polymorphic sites were constituted by two parts of 9 singleton variable sites (two variants) and 38 parsimony informative sites covered 29 two variants, 8 three variants and 1 four variants. The estimation of overall transition/transversion bias (R) was 1.052 when frequencies were observed for four nucleotides (A, 23.30%; T, 21.00%; C, 23.20%; G, 32.50%) (Table 2). The molecular genetic indices analyses showed that 29 transitional sites and 30 transversional sites were estimated for all 48 pigs. Among these substitutions, the rates of different transitions ranged from 9.01 to 15.86 with the highest T-C transition, while the transversional rates ranged from 5.04 to 7.86 with the highest of T-G and C-G transversion (Table 2).

Codon usage in black Dahe pig

The overall codon usage analysis in coding sequences of 48 black Dahe pig showed that the average G + C content at the third synonymous variable coding positions was 0.78. The amino acid composition, corresponding codon and their frequency and the relative synonymous codon usage were calculated for all pig individuals (Table 3). The analysis of amino acids composition and their frequency indicated that amino acid compositional rate ranged from 0.00 to 11.27, and Phe obtained the highest values followed by lesser adjacent rate of Leu, Glu, Arg and Gly respectively (Table 3). There was clear heterogeneity of codon usage among different amino acid. Firstly, estimated codon compositional frequencies obtained a range from 0.00 to 6.30, while the highest rate in codon UUC and the lowest on CUU, GUU, GCA, UAA, UAG and GAU. Secondly, codon UUC was about six fold as frequent as UUU in the preferred Phe amino acid, while the codon CCA and UGU obtained a compositional rate of 1.00 in the secondly non-preferred amino acid of Pro and Cys. The trend of codon usage pattern indicated a significant biased codon usage because of majority of relative synonymous codon usage (RSCU) values near the range of 1.00 to 2.00, but obvious exceptions were observed for four comparative large RSCU in codons of GUG, CCA, CGG and AGC (Table 3). Correlative analyses showed that there was a significant positive correlation between the codon frequency and RSCU values (R^2 = 0.442) (Figure 1).

DISCUSSION

We have presented evidence that the exon 2 of SLA-DRB of black Dahe pig obtained the variable rate of 23.50%, which was higher than the values of 13.70% analyzed for cattles. The results also showed that the parsimony informative sites not only included two variants, three variants and four variants, but even accounted for nearly 80.85% of the 47 polymorphic sites. All the aforementioned results together indicated that there was higher polymorphism in the exon 2 of SLA-DRB gene of black Dahe pig than in other class II genes, which could be consistent with the supported viewpoint of previous studies (Hughes and Nei, 1989; Kloch et al., 2010; Spurgin and Richardson, 2011). In this study, a better explanation for this high polymorphism could be the over dominance selection that emphasized enhancing the rate of amino acid substitution and increasing the heterozygosity and persistence time of polymorphic alleles as compared with those of neutral alleles (Hughes and Nei, 1989). The transitions of T-C and C-T pyrimidine achieved the highest values of 15.86 and 14.00 respectively among all transitional substitutions in this study (Table 2). These may induce enhancing of amino acid substitution rate (Table 3) instead of decreasing of disadvantageous selection, because the selection from advantageous mutations could reduce the level of nucleotide polymorphism (Akashi, 1994; Hughes and Nei, 1989). This important genetic information of high

Table 3. The amino acid composition and relative synonymous codon usage.

Amino acid	AF	Codon	F	RSCU	Amino acid	AF	Codon	F	RSCU
Phe	11.27	UUU	1.00	0.27	Ser	4.74	UCU	0	0
		UUC	6.30	1.73			UCC	1.10	2.21
							UCA	0	0
Leu	10.87	UUA	1.40	1.21			UCG	0	0
		UUG	1.80	1.57					
					Pro	1.59	CCU	0	0
Leu		CUU	0.10	0.11			CCC	0	0.13
		CUC	1.20	1.05			CCA	1.00	3.88
		CUA	0	0			CCG	0	0
		CUG	2.40	2.06					
					Thr	4.63	ACU	0	0
Ile	0.00	AUU	0	0			ACC	1.40	1.89
		AUC	0	0			ACA	1.00	1.38
		AUA	0	0			ACG	0.50	0.73
Met	0.05	AUG	0	1.00	Ala	3.89	GCU	0	0.05
Val	7.03	GUU	0.10	0.06			GCC	1.00	1.59
		GUC	0	0.03			GCA	0.10	0.1
		GUA	0	0			GCG	1.40	2.26
		GUG	4.50	3.91					
					Cys	1.59	UGU	1.00	2.00
Tyr	5.08	UAU	1.20	0.71			UGC	0	0
		UAC	2.10	1.29					
Ter		UAA	0.10	1.50	ter		UGA	0	0
Ter		UAG	0.10	1.50	Trp	0.00	UGG	0	0
His	3.63	CAU	1.70	1.42	Arg	9.32	CGU	0.30	0.32
		CAC	0.70	0.58			CGC	1.00	1.03
							CGA	0	0
Gb		CAA	0.30	0.41			CGG	2.90	2.86
		CAG	1.30	1.59					
					Ser	4.74	AGU	0	0
Asn	4.74	AAU	0.30	0.21			AGC	1.90	3.79
		AAC	2.70	1.79					
					Arg	9.32	AGA	0.60	0.58
Lys	4.59	AAA	0.80	0.54			AGG	1.20	1.22
		AAG	2.20	1.46					
					Gly	8.28	GGU	0	0
Asp	6.03	GAU	0.10	0.03			GGC	1.00	0.72
		GAC	3.80	1.97			GGA	2.00	1.52
							GGG	2.40	1.76
Glu	10.22	GAA	0.90	0.27					
		GAG	5.70	1.73					

AF, Average amino acid frequency; F, codon frequency; RSCU, relative synonymous codon usage.

polymorphism in exon 2 of SLA-DRB of black Dahe pig could provide us a good molecular marker for studying disease resistance breeding and correlative analyses of economic characters.

The results also supported that there was a significant nucleotide base bias in the exon 2 of SLA-DRB gene of black Dahe pig. The content of G + C in the 200 sites was nearly 55.70%, and increased to 78% at the third synonymous variable coding positions (Table 2). This phenomenon of nucleotide bias could still influence codon usage and frequency of acid amino substitution, such as majority of RSCU values near the range of 1.00

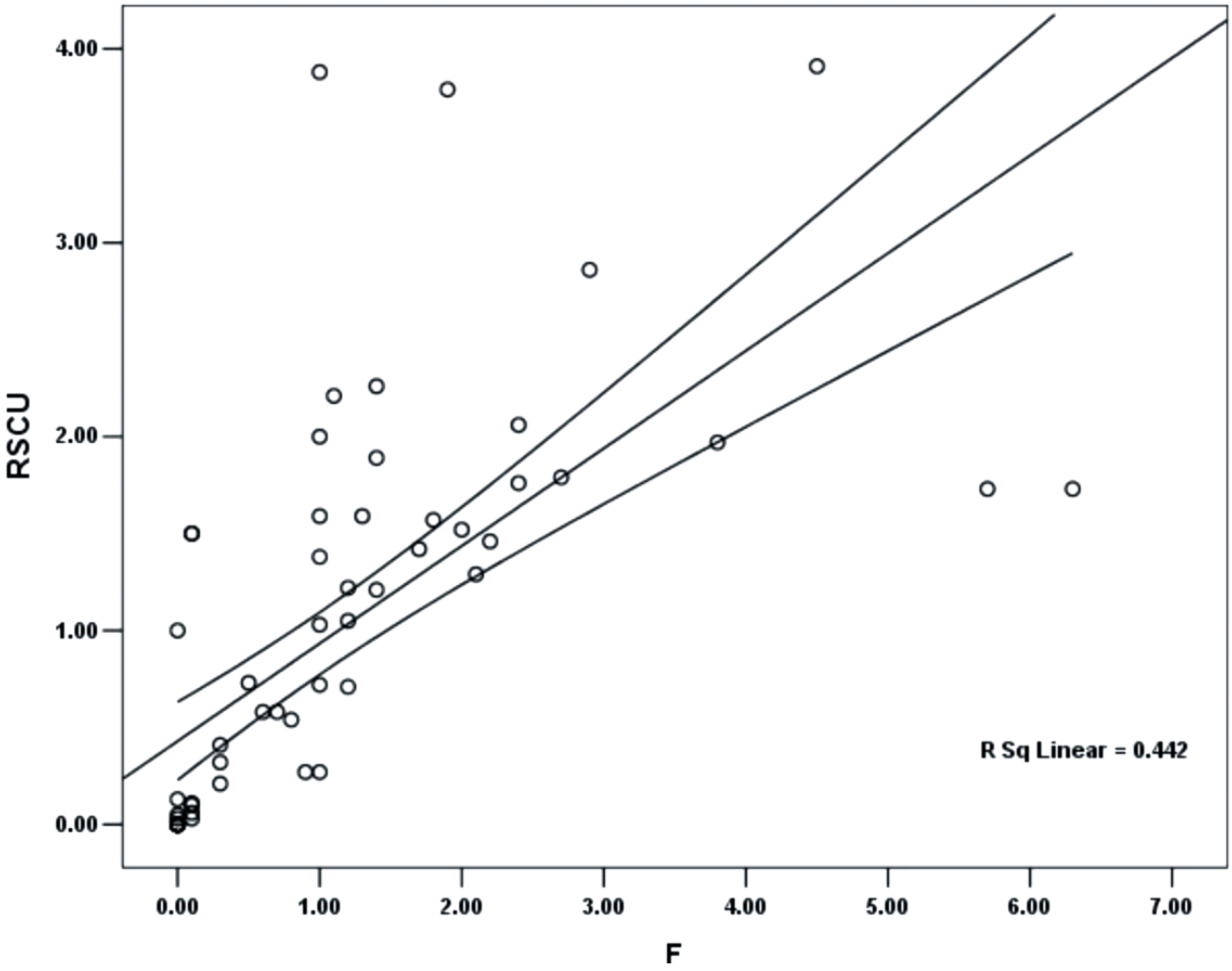

Figure 1. The correlation between codon frequency and relative synonymous codon usage. The above and below lines indicate the 95% confidence interval.

to 2.00, and the amino acid compositional rate ranged from 0.00 to 11.27 (Table 3). This significant bias of RSCU could be dominated by base compositional bias and also be associated with gene expression level (Gupta and Ghosh, 2001). Such codon usage biases may be result from natural selection acting on silent changes in DNA, or both mutational biases and natural selection (Bulmer, 1991; Akashi, 1994; Duret and Mouchiroud, 1999; Grocock and Sharp, 2002). The selective advantage may depend on increasing the translation efficiency, especially during periods of competitive exponential growth (Duret and Mouchiroud, 1999; Grocock and Sharp, 2002).

In conclusion, the significant polymorphism, high GC_3 content and preferred codon usage could indicate that the exon 2 of SLA-DRB gene of black Dahe pig must be an important genetic marker of chromosome for studies in genetic diversity, population genetics, breeding reproduction investigation and exploring new functional genes for disease resistance. The acquired sequences information of SLA-DRB could be useful for the researches on the relationship between the SLA haplotypes and the economic characters of black Dahe pigs, despite fewer samples of pigs used in this study.

ACKNOWLEDGEMENTS

This work was supported by the NSFC (30970366 to T-Z Zhang), the Applied Fundamental Research Project of Yunnan Provincial Science and Technology Department (2010ZC149 to L-Z Tang), and the Key Project of Qujing Normal University Scientific Research Fund (2009ZD004 to L-Z Tang).

REFERENCES

Akashi H (1994). Synonymous codon usage in Drosophila melanogaster: Natural selection and translational accuracy. Genetics, 136: 927-935.

Benacerraf B (1981). Role of MHC gene products in immune regulation. Science, 212: 1229-1238.
Bulmer M (1991). The selection-mutation-drift theory of synonymous codon usage. Genetics, 129: 897-907.
Duret L, Mouchiroud D (1999). Expression pattern and, surprisingly, gene lenth shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. Proc. Natl. Acad. Sci. USA. 96: 4482-4487.
Gorer PA (1937). The genetic and antigenic basis of tumour transplantation. J. Pathol. Bacteriol. 44: 691-697.
Grocock RJ, Sharp PM (2002). Synonymous codon usage in Pseudomonas aeruginosa PA01. Gene, 289: 131-139.
Gu X, Nei M (1999). Locus-specificity of polymorphic alleles and evolution by birth-and-death process in mammalian MHC genes**.** Mol. Biol. Evol. 16: 147-156.
Gupta SK, Ghosh TC (2001). Gene expressivity is the main factor in dictating the codon usage variation among the genes in *Pseudomonas aeruginosa*. Gene, 273: 63-70.
Hughes AL, Nei M (1989). Nucleotide substitution at major histocompatibility complex class II loci: Evidence for overdominant selection. Proc. Natl. Acad. Sci. USA. 86: 958-962.
Huo JL, Huo HL, Miao YW, Li FQ, Liu LX, Wu GM, Ouyang YN, Qian K (2009). Genetic diversity of 76 STR loci in the Dahe pig. Zool. Res. 30: 105-108.
Kloch A, Babik W, Bajer A, Sinski E, Radwan J (2010). Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole *Myodes glareolus*. Mol. Ecol. 19: 255-265.
Lee JH, Simond D, Hawthorne WJ, Walters SN, Patel AT, Smith DM, O'connell PJ, Moran C (2005). Characterization of the swine major histocompatibility complex alleles at eight loci in Westran pigs. Xenotransplantation, 12: 303-307.
Rabin M, Fries R, Singer D, Ruddle FH (1985). Assignment of the porcine major histocompatibility complex to chromosome 7 by in situ hybridization. Cytog. C. Gen. 39: 206-209.
Rozas J, Sanchez-Delbarris JC, Messeguer X, Rozas R (2003). DnsSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19: 2496-2497.
Shia YC, Bradshaw M, Rutherford MS, Lewin HA, Schook LB (1995). PCR-based genotyping for characterization of SLA-DQB and SLA-DRB alleles in domestic pigs. Anim. Genet. 26: 91-99.
Spurgin LG, Richardson DS (2011). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc. Natl. Acad. Sci. USA. 277: 979-988.
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 28: 2731-2739.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The Clustal_X windows interface: flexible strategies for multiple alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882.
Wu SL, Ju HP, Sun PX, Huang J, Bao WB, Hua JD, Huang XG, Chen GH (2007). Polymorphism of exon 2 of SLA-DQB and SLA-DRB genes and its relationship for productive performance in Sutai pigs. J. Agric. Biotech. 15: 606-611.
Xu RH, Wang K, Sun DX, Ding XD, Liu PQ, Zhang Q (2005). Sequences and their polymorphisms of Swine MHC-DQB, DRB proximal promoter region. Acta Genet. Sin. 32: 282-288.