Arch. Anim. Breed., 59, 275–283, 2016 www.arch-anim-breed.net/59/275/2016/ doi:10.5194/aab-59-275-2016 © Author(s) 2016. CC Attribution 3.0 License.





Detection of one large insertion/deletion (indel) and two novel SNPs within the *SPEF2* gene and their associations with male piglet reproduction traits

Rui Chen^{1,2}, Shuai Yu², Fa Ren², Xiao Yan Lv³, and Chuan Ying Pan²

¹Innovation Experimental College, Northwest A&F University, Yangling 712100, China
²College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China
³National Swine Foundation Seed Farm of Ankang Yangchen Modern Agriculture Group Co. Ltd, Ankang 725000, China

Correspondence to: Chuan Ying Pan (panyu1980@126.com) and Xiao Yan Lv (lxy1018lxy1018@163.com)

Received: 26 January 2016 - Revised: 29 April 2016 - Accepted: 17 May 2016 - Published: 15 June 2016

Abstract. The sperm flagella 2 (*SPEF2*) gene is essential for normal sperm tail development and male fertility. To fully characterize the structure of the mutation and to further study the function of the pig *SPEF2* gene, we explored the insertion/deletion (indel) and novel single-nucleotide polymorphisms (SNPs) within the pig *SPEF2* gene, and tested their associations with the testicular sizes in male Large White (LW) and Landrace (LD) pigs from China. Herein, a large insertion located at the *SPEF2* gene in chromosome 16 was found, and two alleles of "T" (insertion) and "D" (deletion) were designated. Allele "D" was dominant in all analyzed pigs. Two novel SNPs (namely (NC_010458) g.19642G > A, resulting in *AfaI* aCRS PCR–PFLP, and g.19886C > G, resulting in *Eco*RI aCRS PCR–PFLP) were found in LW and LD pigs. Association testing revealed that g.19886C > G was significantly associated with the testis long circumference (TLC) in LW pigs (*P* < 0.05), suggesting that this SNP would be the DNA marker for the marker-assisted selection (MAS) in reproduction traits. This preliminary result indicates that the pig *SPEF2* gene had significant effects on male reproduction traits. These findings could not only extend the spectrum of genetic variations in the pig *SPEF2* gene but also contribute to implementing MAS in genetics and breeding in pigs.

1 Introduction

The pig (*Sus scrofa*) is an economically important livestock animal for meat production (Uddin et al., 2011; Park et al., 2015). It is also being increasingly exploited as an ideal animal model in human disease due to its high similarity with regard to human genome size and physiological characteristics (Swindle et al., 2012; Wang et al., 2015). Large White (LW) and Landrace (LD) pigs are the most popular breeds in many countries, including China (Bergfelder-Drüing., 2015), yet few studies on their reproduction traits have been published. Reproduction traits of livestock are important because of the major roles they play in the economic success of production (Rothschild et al., 1996). Moreover, as is well known, reproduction traits have been determined for both females and males (Beerda et al., 2008; Pausch et al., 2014). To date, numerous studies on reproduction traits of female pigs have been published; however, little research based on male sterility of pigs has been done. Therefore, it is necessary to study the reproductive performance of male pigs.

The sperm flagella 2 (*SPEF2*, also known as *KPL2*) gene is essential for normal sperm tail development and male fertility (Sironen et al., 2002; Guo et al., 2013). In rats, the *SPEF2* gene was found to be stage-specific and intensive in spermatocytes and round spermatids in the seminiferous tubules (Ostrowski et al., 1999). In mature murine sperm, *SPEF2* was present in the distal part of the sperm tail mid-piece, sertoli cells and germ cells. Further research has showed that the *SPEF2* gene had high expression in the tail of elongating spermatids and in the tail of mouse sperm (Sironen et al., 2010). The loss of *SPEF2* function in mice resulted in spermatogenesis defects and primary ciliary dyskinesia, which caused a decline in elongating spermatids and faults in the formation of sperm tail (Sironen et al., 2011). In cattle, alternative splicing (AS) has been suggested to be involved in the regulation of *SPEF2* expression in the testes and sperm, and it was one of the determinants of sperm motility during bull spermatogenesis (Guo et al., 2013). In diploid and triploid cyprinid fish, RNA-seq was used to examine fertility-related molecular mechanisms and found that *SPEF2* was expressed at higher levels in the testis of diploid fish than in those of triploids, which indicated that there is a positive correlation between the level of *SPEF2* expression and male fertility (Xu et al., 2015). In pigs, little information on the *SPEF2* gene has been reported, and no reports on genetic variations have been published except regarding the large insertion/deletion (indel).

At the end of 1990s, a reproductive problem within Finnish Yorkshire (Large White) boars was detected when 12 Yorkshire breed boars were found to be affected with the "short-tail" defect, which is associated with 100 % akinesia of sperm (Andersson et al., 2000). In 1998, nine new cases of boars with "short-tail" spermatozoa were detected (Sironen et al., 2002). The causal mutation for this defect was a recent large insertion in intron 30 within the *SPEF2* gene in chromosome 16 (Sironen et al., 2002, 2006). In 2001, the frequency of the mutation was already 23 % (Sironen et al., 2006). It was reported in 2011 that the insertion was associated with increased litter size in the Finnish Yorkshire population (Sironen et al., 2011). Although considerable research has been devoted to the Finnish Yorkshire, no attention has been paid to LW and LD pigs in China.

As one of the important genes related to breeding, the pig *SPEF2* gene has been used as a candidate gene to study the relationships between the indel and single-nucleotide polymorphisms (SNPs) with reproduction traits. Therefore, the aim of this study was to detect one large insertion and novel SNPs within the *SPEF2* gene as well as their associations with male piglet reproduction traits in LW and LD pigs from China. The data could not only extend the spectrum of genetic variations of the pig *SPEF2* gene but also contribute to implementing marker-assisted selection (MAS) in genetics and breeding in pigs.

2 Materials and methods

Experimental animals and procedures used in this study were approved by the Faculty Animal Policy and Welfare Committee of Northwest A&F University (FAPWC-NWAFU) under contract. Furthermore, the care and use of experimental animals completely adhered to the local animal welfare laws, guidelines, and policies.

2.1 Animals and data collection

All testis samples were obtained from 442 male piglets belonging to two breeds (LW and LD) which were housed in the national swine breeding farm, Ankang, Shaanxi, China. LW pigs were of three different ages, among which 75.08 % (n = 250), 6.30 % (n = 21) and 18.62 % (n = 62) were 15, 35 and 40 days old, respectively. LD pigs were of two different ages, among which 9.17 % (n = 10) and 90.83 % (n = 99) were 15 and 40 days old, respectively. Data about testis weight (TW), testis long circumference (TLC) and testis short girth (TSG) were derived from testicular tissues, which were used for association evaluation analysis. All tissues were immediately frozen in liquid nitrogen and stored at -80 °C.

2.2 DNA isolation and primer design

Genomic DNA from 442 samples was isolated following the procedure described by Lan et al. (2007). The quality of genomic DNA was assayed by Nanodrop 1000, and the working concentration of each DNA samples was diluted to $50 \text{ ng }\mu\text{L}^{-1}$ (Jia et al., 2015). Previous data showed that a large insertion in the pig *SPEF2* gene had been identified (Sironen et al., 2012). In order to determine the nature of the insertion, the P1 primer within exon 30 (Table 1) was used in genotyping assays to test for the insertion (Sironen et al., 2012). Also, several other pairs of primers (P2–P7, Table 1) were designed to amplify specific parts of the pig *SPEF2* gene using Primer Premier software (version 5.0) based on the pig *SPEF2* gene sequence (GenBank accession no. NC_010458).

2.3 PCR amplification and DNA sequencing

A total of 50 DNA samples from LW and LD pigs were randomly selected in order to construct genomic DNA pools. Firstly, the genomic DNA pools were used as templates for PCR amplification and explored genetic variation in the pig SPEF2 gene. Then, each individual was assessed to detect whether there was a variation. PCR reactions were performed with touch-down PCR in a 25 µL volume containing 50 ng of genomic DNA, 0.4 µM of two of each primer, 1 × buffer (including 1.5 mM of MgCl₂, 200 µM of dNTPs and 0.625 units of Taq DNA polymerase; MBI, Vilnius, Lithuania) (Kumchoo et al., 2015). The touch-down PCR protocol was as follows: initial denaturation for 5 min at 95 °C, followed by 18 cycles of denaturation for 30s at 94 °C; annealing for 30s at 68 °C (with a decrease of 1 °C per cycle); extension for 1-3 min at 72 °C; another 23 cycles of 30 s at 94 °C, 30 s at 50 °C, and 2 min at 72 °C; and a final extension for 10min at 72 °C, with subsequent cooling to 4 °C. The PCR products were analyzed by means of agarose gel electrophoresis. After that, the PCR products were sent to sequencing when they had a single objective band of each pair of primers.

 Table 1. PCR primer sequences of the pig SPEF2 gene for amplification.

Loci	Primer sequences (from $5'$ to $3'$)	Tm (°C)	Sizes (bp)	Detection methods
P1*	F1: GGCAATATCAAGGTCTTTCCA			
	R1(in): GGCAATATCAAGGTCTTTCCA	TD	421	
	R1(out): GCAGGAGAGGAGAATGACCA	TD	354	cited by Sironen et al. (2012)
P2	F2: GAGCGTGTTCAGCGGACTT (nt 12631-49)	58.3	1007	Pool DNA sequencing
	R2: GGTGGGCTTGCCTGCTAT (nt 13620-37)			
P3	F3: AGGTCAGGGTGTCTTATGC (nt 19263-81)	51.9	1304	Pool DNA sequencing
	R3: ATGGGTTGGTAGGTTGCT (nt 20549-66)			
P4	F4: AATGAGGGTCTGGAGGAGTT (nt 127061-80)	60.0	2137	Pool DNA sequencing
	R4: GTGTAGGTTGCTGATGTGGC (nt 129178-97)			
P5	F5: AGGGAAGACAGCCCGAATG (nt 158740-58)	59.0	1873	Pool DNA sequencing
	R5: ACCACAGCCACAGCAAGG (nt 160595-612)			
P6	F6: GGCAGCAACCATTAGACCC (nt 170882-900)	57.7	1108	Pool DNA sequencing
	R6: CTTCCACCCTTACCACCAA (nt 171971-89)			
P7	F7: GGAAGCCCAGGATTCAAGT (nt 182191-209)	56.3	1235	Pool DNA sequencing
	R7: GGAGGGTAATGTTTGCGGTAG (nt 183405-250)			
P8	F8:AAAGAAATGGACATTTAAA[G]T (nt 19621-41)	51.3	217	AfaI PCR-RFLP
	R8:AGTCATGTAGATACATTGAGATTTT (nt 19813-37)			(GG: 217 bp; AA: 196 bp;
				GA: 217 + 196 bp)
P9	F9:ATATTTAAAAGAATATAGGA[G]AATT (nt 19861-85)	50.9	245	EcoRI PCR RFLP
	R9:TGTATAATTGATATAAGAGTTTCAT (nt 20081-105)			(GG: 245 bp; CC: 224 bp;
				GC: 245 + 224 bp)

[N] shows a mismatch of forward primer for creating a restriction site; F: forward; R: reverse; TD: touch-down PCR protocol.

2.4 Genotyping of the large indel locus and novel SNPs within the pig SPEF2 gene

For large indel locus, according to Sironen et al. (2012), P1 primer was used to test for the insertion. The PCR conditions were the same as touch-down PCR protocol except extension for 30 s at $72 \text{ }^{\circ}\text{C}$.

For SNP locus, the touch-down PCR products sequencing and Blastn analyses were carried out to scan the genetic variants in P2–P7 loci within the pig *SPEF2* gene in analyzed breeds. In this work, two novel SNPs were detected, namely g.19642G > A and g.19886G > C. Since the two SNPs could not be recognized by the natural restriction endonuclease, two pairs of primers P8 (building forced *AfaI*) and P9 (building forced *Eco*RI) were designed to further analyze these SNPs with regard to the *AfaI* and *Eco*RI aCRS PCR–RFLP, respectively (Table 1).

For the g.19642G > A (P8 primer) locus, a new restriction endonuclease *Afa*I site (GTAC) was established by changing the forward primer actual nucleotide "T" to "G" at the g.19640 locus. Then the SNP of g.19642 A with induced point mutation g.19640 G could be genotyped by *Afa*I aCRS PCR–RFLP, but g.19642 G could not.

For the g.19886G>C (P9 primer) locus, a new restriction endonuclease EcoRI site (GAATTC) was established by changing the forward primer actual nucleotide "T" to "G" at the g.19881 locus. Then the SNP of g.19886 C with induced

point mutation g.19881 G could be genotyped by *Eco*RI aCRS PCR–RFLP, but g.19886 G could not.

For the above two SNP loci, $8 \mu L$ of PCR amplification was digested with 3 U A f a I and E co R I, at $37 \degree C$ for 12 h. The digested products were detected by electrophoresis of 3.5 % agarose gel stained with ethidium bromide.

2.5 Statistical analysis

Using PopGene version 1.3.1 (Chen et al., 2013), population parameters such as gene heterozygosity (He), effective allele numbers (N_e) and polymorphism information content (PIC) were calculated (Jia et al., 2015). Genotypic frequencies, allelic frequencies and Hardy-Weinberg equilibrium (HWE) were analyzed using the SHEsis program (http: //analysis.bio-x.cn) (Li et al., 2009), and linkage disequilibrium and haplotypes across two SNPs and the insertion locus were estimated in pig breeds by means of the partitionligation-combination-subdivision expectation maximization algorithm (Wang et al., 2013). Moreover, based on the correlation coefficients (r^2) , the pattern of pairwise linkage disequilibrium between the indel and SNPs was estimated and visualization of linkage disequilibrium measures was determined. Association tests of the polymorphism with three reproduction traits (testis weight (TW), testis long circumference (TLC) and testis short girth (TSG)) were conducted for three different growth periods (15, 35 and 40 days) in LW pigs, and two different growth periods (15 and 40 days)



Figure 1. Structure mode, sequence chromas and electrophoresis pattern of the pig *SPEF2* gene. (a) Position of primers on *SPEF2* gene sequences, (b) sequence chromas of two novel SNP loci of the pig *SPEF2* gene, and (c) electrophoresis pattern of two novel SNPs and an indel of the pig *SPEF2* gene.

in LD pigs. These association analyses were performed using the analysis of variance (ANOVA) procedure in SPSS (version 18.0) when data agreed with the characteristic of normality and homogeneity of variances. If they did not agree, the nonparametric test (Kruskal–Wallis) was conducted in SPSS (18.0). The ANOVA applied the general linear model (GLM), and the reduced linear model was as follows: $Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$, where Y_{ijk} is the observation of the reproduction traits (e.g., testis weight), α_i is the fixed effect of age, β_j is the fixed effect of genotype or combined genotype, and ε_{ijk} is the random residual error (Yang et al., 2015).

3 Result

3.1 Genetic variant identification of the *SPEF2* gene in pigs

In this study, in order to determine the nature of the insertion, the insertion was cloned and sequenced. An illustration of the insertion region is present in Fig. 1. Two genotypes (DD and ID) were identified, showed one band (354 bp) and two bands (354 and 421 bp), respectively. The pool sequencing and alignment analysis identified two novel SNPs, namely g.19642G > A and g.19886G > C, within the pig *SPEF2* gene in LW and LD pigs (Fig. 1).

Loci Breeds	Sizes	Genoty	pic frequ	uencies	Allelic fro	equencies	HWE ^a P values	Popula Ho ^b	tion para He ^c	meters $N_{\rm e}^{\rm d}$	PIC ^e
g.19642G>A	Ν	AA	AG	GG	А	G					
LD LW	72 370	0 0	0.429 0.419	0.571 0.581	0.214 0.209	0.786 0.791	>0.05 <0.05	0.663 0.669	0.337 0.331	1.508 1.495	0.280 0.276
g.19886G>C	Ν	CC	CG	GG	С	G					
LD LW	72 370	0.508 0.384	0.369 0.488	0.123 0.128	0.692 0.628	0.308 0.372	> 0.05 > 0.05	0.574 0.533	0.426 0.467	1.741 1.877	0.335 0.358
indel	Ν	II	ID	DD	Insertion	Deletion					
LD LW FY sows* FY boars*	72 370 357 491	0 0 0 0	0.224 0.213 0.401 0.179	0.776 0.787 0.599 0.821	0.112 0.106 0.200 0.09	0.888 0.894 0.800 0.91	>0.05 >0.05	0.810 0.801 0.680 0.837	0.190 0.199 0.320 0.163	1.235 1.249 1.471 1.195	0.172 0.179 0.269 0.150

Table 2. Genotypic, allelic frequencies and population indexes for an indel and two SNPs of the pig SPEF2 gene.

^a HWE: Hardy–Weinberg equilibrium; ^b Ho: homozygosity; ^c He: heterozygosity; ^d N_e: effective allele numbers; ^e PIC: polymorphism information content; LW: Large White pigs; LD: Landrace pigs; indel: insertion/deletion; II: insertion/insertion; DD: deletion/deletion; ID: insertion/deletion; *SPEF2*: sperm flagella 2; *: Sironen et al. (2012); FY: Finnish Yorkshire.

The g.19642G > A locus was genotyped by using the *AfaI* aCRS PCR–RFLP assay. At the *AfaI* locus, genotype GG demonstrated one band (217 bp), genotype AA showed two bands (196 and 21 bp), and genotype CG generated three bands (217, 196 and 21 bp) (Fig. 1). Therefore, two fragments (217 and 196 bp) were utilized to determine the different genotypes.

The g.19886G>C locus was genotyped by using the *Eco*RI aCRS PCR–RFLP assay. At the *Eco*RI locus, genotype GG demonstrated one band (245 bp), genotype CC showed two bands (224 and 21 bp), and genotype AG generated three bands (245, 224 and 21 bp) (Fig. 1). Therefore, two fragments (245 and 224 bp) were utilized to determine the different genotypes.

3.2 Genotype, allele frequency and genetic diversity of the *SPEF2* gene in pigs

The sample sizes, genotypic frequencies, allelic frequencies, homozygosity (Ho), heterozygosity (He), effective allele numbers (N_e), and polymorphism information content (PIC) of LW and LD pigs are shown in Table 2. The results indicate that these loci were polymorphic in LW and LD pigs. The classification of PIC values demonstrated that all SNPs were close to medium genetic diversity, though the indel locus was close to low genetic diversity. The χ^2 test indicated that these loci were at HWE (P > 0.05), except for the g.19642G > A locus in LW pigs. The frequencies of genotypes and main alleles were different at different SNP loci and indel locus in two pig breeds. For example, as shown in Table 2, the frequencies of the two alleles of g.19642G > A were similar in LW and LD pigs.

The insertion frequencies of the *SPEF2* gene in LW and LD pigs were 10.6 and 11.2 %, respectively. Two genotypes (GG and AG) were shown in *AfaI* aCRS PCR–RFLP analysis in the analyzed population (Table 2). The frequencies of genotypes GG and AG were 0.581 and 0.419 in LW pigs and 0.571 and 0.429 in LD pigs. Three genotypes (CC, CG and GG) were shown in the *Eco*RI aCRS PCR–RFLP analysis in the analyzed population (Table 2). The frequencies of genotypes CC, CG and GG were 0.384, 0.488 and 0.128 in LW pigs and 0.508, 0.369 and 0.123 in LD pigs.

3.3 Haplotype structure and linkage disequilibrium analysis of the *SPEF2* gene in pigs

The haplotype frequencies for indel locus and two SNP loci within the pig *SPEF2* gene are shown in Table 3. Seven haplotypes were found in LW pigs, while three haplotypes were found in LD pigs (Table 3), which was lower than excepted (n = 8). The three haplotypes (hap 1, hap 5 and hap 6) were simultaneously found in the analyzed populations. Hap 5 was the major haplotype in both LW (48.0%) and LD (75.0%) pigs.

The linkage disequilibrium of indel locus and two SNP loci in the two pig breeds was analyzed. As shown in Table 4 and Fig. 2, the r^2 values of LW pigs were very low but the D' values were moderate. The D' values between g.19642G>A and g.19886G>C, g.19642G>A and indel, and g.19886G>C and indel were 0.601, 0.285 and 0.499, respectively. As shown in Table 4 and Fig. 2, the r^2 values of LD pigs were very low. Most of the D' values were low, except the D' values between g.19642G>A and g.19886G>C and indel, which were 1.000 and 0.474, respectively. Although the r^2 values of LD pigs be-

Different Haplotypes	g.19642G > A - g.19886G > C - indel	Haplotyp LW	e frequencies LD
Hap1	A C D	0.097	0.167
Hap2	A C I	0.067	_
Hap3	A G D	0.036	_
Hap4	A G I	0.001	-
Hap5	G C D	0.480	0.750
Hap6	G G D	0.242	0.083
Hap7	GGI	0.077	-

Table 3. Haplotype frequencies within the SPEF2 gene in pig breeds.

"Hap" represents "haplotype"; "–" means P < 0.01; LW: Large White; LD: Landrace; indel: insertion/deletion.

Table 4. D' and r^2 values of pairwise linkage disequilibrium of the *SPEF2* gene in pig breeds.

Breeds]	LW breed		LD breed			
	g.19642G>A	g.19886G>C	indel	g.19642G>A	g.19886G>C	indel	
g.19642G>A	0.050	0.601	0.285	0.001	1.000	0.000	
g.19886G>C indel	0.052 0.055	0.021	0.499	0.021 0.000	0.014	0.474	

Note: D' and r^2 values for pairwise linkage disequilibrium analysis are shown by the upper and lower triangles in the table, respectively.

tween g.19642G > A and g.19886G > C were low, at the same time, the D' values were high (1.000); therefore, we analyzed the effects of the interaction between g.19642G > A and g.19886G > C of LD pigs with reproduction traits as well as between g.19642G > A and g.19886G > C (0.601) and between g.19886G > C and indel (0.499) of LW pigs.

3.4 Association study of the *SPEF2* gene genetic variations with reproduction traits in pigs

Before analysis, significant differences were found between the pig ages (10, 35 and 40 days) and the TW, TLC and TSG traits in LW and LD pigs (P < 0.01). However, this work focused on the relationship between the different genotypes and reproductive traits of the same breed. Thus, age was regarded as a fixed factor in the linear model. Herein, the association of the indel locus and the two SNP loci with reproduction traits of the specific stage was investigated in the tested breeds. Firstly, the effects of the indel locus and the two SNP loci on reproduction traits were evaluated in both LW and LD pigs. In the g.19886G > C locus, the different genotypes were found to be significantly associated with the TLC trait in LW pigs (P < 0.05) (Table 5), which demonstrated that the genotype GG was superior to CG in LW pigs. TLC was highest in the G-allele homozygotes (6.98 ± 0.23 cm), intermediate in the C-allele homozygotes $(6.34 \pm 0.12 \text{ cm})$, and lowest in the CG heterozygotes $(6.16 \pm 0.15 \text{ cm})$. However, for the indel locus and the g.19642G > A locus, there were no statistically significant differences detected between the

three genotypes and reproduction traits in the LW and LD pigs (P > 0.05) (data not shown).

4 Discussion

The SPEF2 gene participated in spermatogenesis, sperm tail formation and ciliary function (Sironen et al., 2002, 2011; Guo et al., 2013). A reproductive problem within the Finnish Yorkshire boars was first detected at the end of 1990s, referred to as a "short-tail" defect, which is caused by a large insertion within the SPEF2 gene in chromosome 16; this conclusion has been further proved by several experiments (Andersson et al., 2000; Sironen et al., 2002, 2006, 2007). In 2001, it was reported that the frequency had increased to 36 % (Sironen et al., 2002), but after MAS the frequency reduced to 18% in 2005 (Sironen et al., 2007). However, the "short-tail" defect in LW and LD pigs from China remains unknown. So far, there have been no reports describing the relationships between polymorphisms of the SPEF2 gene and testis measurement traits in pigs. Therefore, evaluating the associations between the large insertion and novel SNPs of the pig SPEF2 gene and testis measurement traits was necessary and of interest.

In 2011, it was reported that the insertion of the *SPEF2* gene in chromosome 16 was associated with increased litter size in the Finnish Yorkshire population (Sironen et al., 2011), but no "short-tail" defect homozygous pigs and only a few homozygous pigs were available for analysis in the study. The insertion frequencies of the *SPEF2* gene in the



Figure 2. Linkage disequilibrium plot of the SPEF2 gene in LW and LD pigs. (a) D' of LW, (b) r^2 of LW, (c) D' of LD, and (d) r^2 of LD.

Locus/reproduction traits	Observe	$SM \pm SE$)	P value	
g.19886G>C				
LW	CC	CG	GG	
TLC	6.34 ± 0.12	6.16 ± 0.15^{b}	6.98 ± 0.23^a	P = 0.012

 Table 5. Relationship between the g.19886G > C locus of the pig SPEF2 gene and reproduction traits.

SPEF2: sperm flagella 2; LW: Large White; TLC: testis long circumference; LSM: least-squares means; SE: standard error. The values with different letters (^a and ^b) within the same row differ significantly at P < 0.05.

analyzed 357 Finnish Yorkshire sows and 491 Finnish Yorkshire boars were 20 and 9%, respectively (Sironen et al., 2011). In this study, mutation of the *SPEF2* gene in 370 LW pigs and 72 LD pigs was detected. Two genotypes (DD and ID) were identified. Genotype DD showed one band (354 bp), while genotype ID demonstrated two bands (354 and 421 bp). The results indicated that no homozygous male pigs and few heterozygous were found in the tested breeds. The insertion frequencies in LW and LD pigs were 10.6 and 11.2 %, respectively. Our results are in agreement with the in-

sertion frequency in previous publications, in particular those of Sironen et al. (2011). However, a relatively lower insertion frequency was proved in our study, which was possibly caused by MAS reducing the number of "short-tail" defect pigs. We speculated that the insertion of the *SPEF2* gene in chromosome 16 may influence the reproductive performance of boars, such as the quality of sperm, sperm motility and spermatogenesis, in turn affecting the total number of piglets born in the first parity.

In this study, two novel SNPs of the *SPEF2* gene, which were located at the intron, were first found in pigs. Moreover, *Afa*I aCRS PCR–RFLP and *Eco*RI aCRS PCR–RFLP were first used to detect and confirm the g.19642G > A and g.19886G > C loci, respectively. For the g.19642G > A locus, the LW breed was not at HWE (P<0.05), which was statistically identified by the lower observed number of genotype AA. A possible reason for this was the use of rapid, powerful and effective selection strategies, which might change the allelic balance of this locus.

The haplotype structure of the pig *SPEF2* gene was also detected. Haplotype 5 (Hap5) was verified as the mutual haplotype in these two breeds with a relatively high frequency. Interestingly, the frequency of hap 5 was varied in these two breeds, which was probably caused by breed distinctiveness.

The association of the two novel SNPs and reproduction traits was analyzed. In the g.19886G>C locus, the different genotypes were found to be significantly associated with the testis size in LW pigs (P < 0.05) (Table 5). Moreover, genotype GG was superior to the other genotypes in LW testis long circumference, suggesting that the allele G of the pig SPEFE gene had positive effects on testis measurement traits in this breed. However, there were no significant differences between g.19886G>C and reproduction traits in the LD breed. This phenomenon was possibly due to the different species. Ghorbankhani (2015) reported that testicular circumference influenced the reproduction of Sanjabi growing ram lambs, independent of nutritional status (Ghorbankhan et al., 2015). In some livestock, testicular volume could reflect spermatogenesis (Gouletsou et al., 2008). Therefore, identifying the individual differences of testicular volume is essential for selecting the animals for high sperm production. There are several studies about bulls and rams in which it is reported that testicular morphometry could serve as an indicator of fertility, providing high correlations with sperm production (Rege et al., 2000; Devkota et al., 2008). Hence, according to the close relationship between testis measure traits and reproduction, we suggest that the G allele of the pig SPEF2 gene had positive effects on reproduction traits. Although g.19642G > A and g.19886G > C were intronic mutations, they might also affect alternatively spliced transcripts of mRNA or transcription factor binding, thus affecting phenotype. A classic example of intronic mutation was located at intron 3 of the porcine IGF2 gene. This mutation led to a significant effect in skeletal muscle, which could affect the combination of the IGF2 gene and the nuclear factor, which is probably a repressor, thus affecting the expression of messenger RNA in postnatal muscle (Van Laere et al., 2003).

Briefly, the large indel and two novel SNPs were found, and g.19886G > C was significantly associated with the testis long circumference (TLC) in LW pigs (P < 0.05), which extends the genetic variations spectrum of the pig *SPEF2* gene and contributes to implementing MAS in genetics and breeding in pigs.

Acknowledgements. This work was funded by the Fundamental Research Funds for the Central Universities (NWAFU, no. 2452015145), a China Postdoctoral Science Foundation funded project (no. 2014M560809) and the Shaanxi Provincial Postdoctoral Science Foundation.

Edited by: S. Maak

Reviewed by: J. Pei and three anonymous referees

References

- Andersson, M., Peltoniemi, O. A. T., Makinen, A., Sukura, A., and Rodriguez- Martinez, H.: The hereditary "short tail" sperm defect – a new reproductive problem in Yorkshire Boars, Reprod. Domest. Anim., 35, 59–63, 2000.
- Beerda, B., Wyszynska-Koko, J., Te Pas, M. F., de Wit, A. A., and Veerkamp, R. F.: Expression profiles of genes regulating dairy cow fertility: recent findings, ongoing activities and future possibilities, Animal, 2, 1158–1167, 2008.
- Bergfelder-Drüing, S., Grosse-Brinkhaus, C., Lind, B., Erbe, M., Schellander, K., Simianer, H., and Tholen, E.: A genomewide association study in large white and landrace pig populations for number piglets born alive, PLoS One, 10, e0117468, doi:10.1371/journal.pone.0117468, 2015.
- Chen, F., Shi, J., Luo, Y. Q., Sun, S. Y., and Pu, M.: Genetic Characterization of the Gypsy Moth from China (Lepidoptera, Lymantriidae) Using Inter Simple Sequence Repeats Markers, PLoS One, 8, e73017, doi:10.1371/journal.pone.0073017, 2013.
- Devkota, B., Koseki, T., Matsui, M., Sasaki, M., Kaneko, E., Miyamoto, A., Amaya, Montoya, C., and Miyake, Y.: Relationships among age, body weight, scrotal circumference, semen quality and peripheral testosterone and estradiol concentrations in pubertal and postpubertal Holstein bulls, J. Vet. Med. Sci., 70, 119–121, 2008.
- Ghorbankhani, F., Souri, M., Moeini, M. M., Mirmahmoudi, R.: Effect of nutritional state on semen characteristics, testicular size and serum testosterone concentration in Sanjabi ram lambs during the natural breeding season, Anim. Reprod. Sci., 153, 22–28, 2015.
- Gouletsou, P. G., Galatos, A. D., Leontides, L. S.: Comparison between ultrasonographic and caliper measurements of testicular volume in the dog, Anim. Reprod. Sci., 108, 1–12, 2008.
- Guo, F., Yang, B., Ju, Z. H., Wang, X. G., Qi, C., Zhang, Y., Wang, C. F., Liu, H. D., Feng, M. Y., Chen, Y., Xu, Y. X., Zhong, J. F., and Huang, J. M.: Alternative splicing, promoter methylation, and functional SNPs of sperm flagella 2 gene in testis and mature spermatozoa of Holstein bulls, Reproduction, 147, 241–252, 2013.

- Jia, W. C., Wu, X. F., Li, X. C., Xia, T., Chen, H., Pan, C. Y., Lan, X. Y.: Novel genetic variants associated with mRNA expression of STAT3 gene significantly affected goat growth traits, Small Ruminant Res., 129, 25–36, 2015.
- Kumchoo, T. and Mekchay, S.: Association of non-synonymous SNPs of *OPN* gene with litter size traits in pigs, Arch. Anim. Breed., 58, 317–323, doi:10.5194/aab-58-317-2015, 2015.
- Lan, X. Y., Pan, C. Y., Chen, H., Zhang, C. L., Li, J. Y., Zhao, M., Lei, C. Z., Zhang, A. L., and Zhang, L.: An AluI PCR-RFLP detecting a silent allele at the goat POU1F1 locus and its association with production traits, Small Ruminant Res., 73, 8–12, 2007.
- Li, Z. Q., Zhang, Z., He, Z., Tang, W., Li, T., Zeng, Z., He, L., and Shi, Y. Y.: A partition-ligation-combination- subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (http://analysis.bio-x.cn), Cell Res., 19, 519– 523, 2009.
- Ostrowski, L. E., Andrews, K., Potdar, P., Matsuura, H., Jetten, A., and Nettesheim, P.: Cloning and characterization of *KPL2*, a novel gene induced during ciliogenesis of tracheal epithelial cells, Am. J. Resp. Cell Mol., 20, 675–683, 1999.
- Park, S. J., Kwon, S. G., Hwang, J. H., Park, da. H., Kim, T. W., Kim, C. W.: Selection of appropriate reference genes for RTqPCR analysis in Berkshire, Duroc, Landrace, and Yorkshire pigs, Gene, 558, 152–158, 2015.
- Pausch, H., Kölle, S., Wurmser, C., Schwarzenbacher, H., Emmerling, R., Jansen, S., Trottmann, M., Fuerst, C., Götz, K. U., and Fries, R.: A nonsense mutation in TMEM95 encoding a nondescript transmembrane protein causes idiopathic male subfertility in cattle, PLoS Genet., 10, e1004044, doi:10.1371/journal.pgen.1004044, 2014.
- Rege, J. E., Toe, F., Mukasa-Mugerwa, E., Tembely, S., Anindo, D., Baker, R. L., and Lahlou-Kassi, A.: Reproductive characteristics of ethiopian highland sheep. II. genetic parameters of semen characteristics and their relationships with testicular measurements in ram lambs, Small Ruminant Res., 37, 173–187, 2000.
- Rothschild, M., Jacobson, C., Vaske, D., Tuggle, C., and Wang, L. Z.: The estrogen receptor locus is associated with a major gene influencing litter size in pigs, P. Natl. Acad. Sci. USA, 93, 201– 205, 1996.
- Sironen, A., Andersson, M., Uimari, P., Vilkki, J.: Mapping of an immotile short tail sperm defect in the Finnish Yorkshire on porcine chromosome 16, Mamm. Genome., 13, 45–49, 2002.
- Sironen, A., Thomsen, B., Andersson, M., Ahola, V., Vilkki, J.: An intronic insertion in *KPL2* results in aberrant splicing and causes the immotile short-tail sperm defect in the pig, P. Natl. Acad. Sci. USA, 103, 5006–5011, 2006.

- Sironen, A., Vilkki, J., Bendixen, C., and Thomsen, B.: Infertile Finnish Yorkshire boars carry a full-length LINE-1 retrotransposon within the *KPL2* gene, Mol. Genet. Genomics, 278, 385–391, 2007.
- Sironen, A., Hansen, J., Thomsen, B., Andersson, M., Vilkki, J., Toppari, J., Kotaja, N.: Expression of *SPEF2* during mouse spermatogenesis and identification of IFT20 as an interacting protein, Biol. Reprod., 82, 580–590, 2010.
- Sironen, A., Kotaja, N., Mulhern, H., Wyatt, T. A., Sisson, J. H., Pavlik, J. A., Miiluniemi, M., Fleming, M. D., Lee, L.: Loss of *SPEF2* function in mice results in spermatogenesis defects and primary ciliary dyskinesia, Biol. Reprod., 85, 690–701, 2011.
- Sironen, A., Uimari, P., Iso-Touru, T., and Vilkki, J.: L1 insertion within SPEF2 gene is associated with increased litter size in the Finnish Yorkshire population, J. Anim. Breed. Genet., 129, 92– 97, 2012.
- Swindle, M. M., Makin, A., Herron, A. J., Clubb, F. J., Frazier, K. S.: Swine as models in biomedical research and toxicology testing, Vet. Pathol., 49, 344–356, 2012.
- Uddin, M. J., Cinar, M. U., Tesfaye, D., Looft, C., Tholen, E., and Schellander, K.: Age-related changes in relative expression stability of commonly used housekeeping genes in selected porcine tissues, BMC. Res. Notes, 4, doi:10.1186/1756-0500-4-441, 2011.
- Van Laere, A. S., Nguyen, M., Braunschweig, M., Nezer, C., Collette, C., Moreau, L., Archibald, A. L., Haley, C. S., Buys, N., Tally, M., Andersson, G., Georges, M., and Andersson, L.: A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig, Nature, 425, 832–836, 2003.
- Wang, A. L., Zhang, Y., Li, M. J., Lan, X. Y., Wang, J. Q., and Chen, H.: SNP identification in FBXO32 gene and their associations with growth traits in cattle, Gene, 515, 181–186, 2013.
- Wang, W., Liu, Y., Wang, H., Ding, X., Liu, J., Yu, Y., Zhang, Q.: A genomic variant in IRF9 is associated with serum cytokine levels in pig, Immunogenetics, 68, 67–76, 2015.
- Xu, K., Wen, M., Duan, W., Ren, L., Hu, F. Z., Xiao, J., Wang, J., Tao, M., Zhang, C., Wang, J., Zhou, Y., Zhang, Y., Liu, Y., and Liu, S. J.: Comparative analysis of testis transcriptomes from triploid and fertile diploid cyprinid fish, Bilo. Reprod., 92, doi:10.1095/biolreprod.114.125609, 2015.
- Yang, D. and Liu, Y.: Molecular cloning, sequence identification, polymorphism and association of the porcine *SPATS2L* gene, Arch. Anim. Breed., 58, 445–449, doi:10.5194/aab-58-445-2015, 2015.