

# Detection of four polymorphisms in 5' upstream region of *PNPLA2* gene and their associations with economic traits in pigs

Lihe Dai<sup>1</sup> · Xiaohong Chu<sup>1</sup> · Fuzeng Lu<sup>1</sup> · Ruhai Xu<sup>1</sup>

Received: 5 February 2016 / Accepted: 22 August 2016 / Published online: 26 August 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** As an important triglyceride hydrolase in mammalian cells, patatin-like phospholipase domain-containing 2 (*PNPLA2*) predominantly performs the first step in triglyceride hydrolysis. The objective of this study was to detect and evaluate the effects of mutations in the 5' upstream region of porcine *PNPLA2* gene with fat deposition and carcass traits. Four single nuclear polymorphisms were identified, including g.161969 T>C, g.161962 A>G, g.161953 C>G and g.161904 G>T, and subsequently genotyped in five pure breeds. Three haplotypes were constructed, including H1(CGGT), H2(TACG) and H3(CACT), which were the most abundant haplotypes in Duroc (0.75), Landrace (0.78) and Chinese indigenous breeds (>0.73), respectively. Duroc individuals with the H1H1 diplotype always exhibited the lowest feed conversion ratio (FCR) ( $P < 0.05$ ), while H2H2 had the thickest backfat thickness ( $P < 0.05$ ). Landrace individuals with H2H3 had lower backfat thickness ( $P < 0.05$ ), higher muscle thickness ( $P < 0.05$ ) and estimated lean meat percentage ( $P < 0.05$ ) than those with diplotype H2H2 and H3H3. Luciferase assay indicated pGL3-basic-H2 had the highest activity and pGL3-basic-H1 had the lowest activity in driving reporter gene transcription in HEK293 cells in vitro. In H1 haplotype, two GR binding sites and an ER $\alpha$  binding site were predicted to be introduced. While in H2 and H3, there were other transcriptional factor binding sites predicted in H2 and H3, such as Sp1, AP-2 and CAC-

binding proteins, which were broadly expressed transcription factors and capable of contributing to basal promoter activity. The reduced basal promoter activity of H1 may be due to the lack of inducement for GR and ER $\alpha$  binding sites in HEK293 cells. The identified functional polymorphisms provide new evidence of *PNPLA2* as an important candidate gene for fat deposition and carcass traits in pigs.

**Keywords** *ATGL* · SNP · *Sus scrofa* · Promoter · Dual-luciferase assay

## Introduction

Patatin-like phospholipase domain-containing 2 (*PNPLA2*), also called adipose triglyceride lipase (*ATGL*) [1], the  $\zeta$ -isoform of a calcium-independent phospholipase A2 (*iPLA2* $\zeta$ ) [2] and desnutrin [3], was identified simultaneously by several laboratories in 2004 as a key enzyme for adipose lipid mobilization. It catalyzes the first step in triglyceride hydrolysis in mammalian adipocyte and non-adipocyte cells [4, 5]. Porcine *PNPLA2* gene is located in a region on porcine chromosome 2p17 (SSC2pter) [6], which harbours a quantitative trait locus for the trait of backfat thickness [7].

Previous studies have discovered functional mutations in *PNPLA2* gene. In human, mutations in *PNPLA2* gene were associated with degradation of lipid droplets and disorders related to fat deposition such as type II diabetes [8, 9]. In obesity model pig, our previous research had identified a missense mutation of G/A392 (R/H131) in porcine *PNPLA2* gene and found the polymorphism was significantly associated with several economic traits detected, including subcutaneous fat thickness, viscera adipose tissue weight, lean meat percentage (LMP), loin

✉ Ruhai Xu  
xuruhai@163.com

<sup>1</sup> Key Laboratory of Animal Genetics and Breeding of Zhejiang Province, Institute of Animal Husbandry and Veterinary Science, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, People's Republic of China

eye traits and even rib numbers, where allele A was associated with decrease of fat deposition [10]. As expected, porcine *PNPLA2* mRNA was expressed abundantly in backfat, mildly in heart, muscle and small intestine, and undetectable in lung, stomach, liver, spleen, kidney and ovary.

A variety of transcription factor binding sites were predicted in the upstream region of porcine *PNPLA2* gene, such as CCAAT/enhancer binding protein (C/EBP), cyclic AMP response element-binding protein (CREBP1) and myogenic differentiation 1 (MyoD) binding sites [10]. Variations in the promoter region may affect the allele-specific gene expression patterns by altering the binding sites for original transcription factor [11]. We therefore selected the porcine *PNPLA2* gene as a potential candidate gene for fat deposition and carcass traits, identified the mutations in the promoter region, determined their roles on fat deposition related traits in different pig breed populations by haplotype construction and association analysis, and evaluated their effects on transcriptional activity by dual-luciferase reporter assay.

## Materials and methods

### Animals and traits

To investigate the allele frequency, animals from five different pig breed populations were chosen, including 19 Chinese Shengxian spot (SXS) pigs, 24 Chinese Bamei (BM) pigs, 161 Duroc pigs, 63 Landrace pigs and 51 Yorkshire pigs. The association analyses were carried out in Duroc, Landrace and Yorkshire pigs, which were all backup boars from different sires and litters, and kept at eight Pig Performance Testing Stations in Experimental Farms of Zhejiang Academy of Agricultural Sciences under the same housing and feeding conditions from 30 to 100 kg body weight. The parameters measured by Feed Intake Record Equipment (FIRE, American Osborn) on individual pigs during the test were FCR and test daily gain (TDG). Off-test backfat thicknesses of Duroc, Landrace and Yorkshire pigs were measured between 3rd and 4th last lumbar vertebrae and 7 cm from the midline (BF1), and between 3rd and 4th last rib and 7 cm from the midline (BF2) with an A-mode ultrasound device (SFK-Technology, Piglog 105, Denmark). Muscle depths were also measured between 3rd and 4th last rib and 7 cm from the midline (MD) with the same device. Lean meat percentage is acquired by determining the percentage weight of the whole carcass, however, in live animals, the LMP was estimated by Piglog105 automatically according to the instruction, on the basis of different fat and meat depths (BF1, BF2, and MD).

### PCR amplification and polymorphism screening of upstream sequence of porcine *PNPLA2* gene

Genomic DNA was extracted from whole blood samples or ear tissues according to standard phenol/chloroform procedure. According to the working draft sequence of *Sus scrofa* chromosome 2 clone CH242-144M19 (GenBank: CU928976.18), oligonucleotide primers P1F and P1R were designed to amplify and verify the upstream sequence of porcine *PNPLA2* gene (Table 1).

PCR reaction systems were prepared in a final volume of 50  $\mu$ l, containing 25  $\mu$ l 2 $\times$  PrimeSTAR<sup>TM</sup> GC buffer (Takara), 0.5  $\mu$ l PrimeSTAR<sup>TM</sup> HS DNA polymerase (2.5 U/ $\mu$ l, Takara), 4  $\mu$ l dNTPs (2.5 mM), 1  $\mu$ l each primer (10  $\mu$ M) and 100 ng genomic DNA. The touchdown PCR program was set as 10 s at 98  $^{\circ}$ C, followed by 13 cycles of 10 s at 98  $^{\circ}$ C, 10 s at 60–48  $^{\circ}$ C (annealing temperature was decreased 1  $^{\circ}$ C after each cycle), and 3.5 min at 72  $^{\circ}$ C. This touchdown step was followed by 28 cycles of 10 s at 98  $^{\circ}$ C, 10 s at 50  $^{\circ}$ C, and 3.5 min at 72  $^{\circ}$ C. After being checked by 1.5 % agarose gel electrophoresis, the PCR products were purified using QIAquick Gel Extraction Kit (Qiagen), and subjected to direct sequencing by an ABI 3100 Avant DNA sequencer (ABI).

In order to explore the genetic variations of the obtained 5' flanking sequence of porcine *PNPLA2* gene in the tested groups, 20 individuals (five individuals per breed) were sequenced and multiple alignments were analyzed using ClustalW (<http://www.ebi.ac.uk/clustalw/index.html>). The result indicated four single nucleotide polymorphisms (SNPs) in the fragment.

### Bioinformatics analysis of the four SNPs

The obtained sequence covering the four SNPs was subjected to transcription factor prediction. The transcriptional binding sites were predicted using TESS (<http://www.gene-regulation.com/pub/databases.html>), SignalScan (<http://www.bimas.cit.nih.gov/molbio/signal/>) and MatInspector (Genomatix Software GmbH).

### Polymorphism genotyping

In order to perform genotyping, primers P2F and P2R (Table 1) were designed to amplify the fragment covering the polymorphic sites in 5' flanking sequence of porcine *PNPLA2* gene. Each reaction mixture (50  $\mu$ l) contained 1.5  $\mu$ l DNA, 10  $\mu$ M primers and 25  $\mu$ l GoTaq Green Master Mix (Promega). The PCR program was set as 2 min at 95  $^{\circ}$ C, 33 cycles of 30 s at 95  $^{\circ}$ C, 10 s at 58.5  $^{\circ}$ C, 50 s at 72 $^{\circ}$ C, and then followed by 5 min at 72  $^{\circ}$ C. PCR products were purified and subjected to direct sequencing by an ABI 3100 Avant DNA sequencer (ABI). Genotypes

**Table 1** List of primers for amplification and plasmid construction

Primer name	Primer sequence (5'-3')	Length (bp)	Usage
P1F	GGGCAGATGTGGAAATGGTTG	3107	Forward primer to amplify upstream sequence of porcine <i>PNPLA2</i> gene
P1R	TAGATGTGCTTGGCGTTGGC		Reverse primer to amplify upstream sequence of porcine <i>PNPLA2</i> gene
P2F	AGGTCCTGCTCTGGTTTCTG	608	Forward primer for SNP genotyping
P2R	CCTCTAAGGGTCTGCCTACAA		Reverse primer for SNP genotyping
P3F	AAAGAGCTCCTGCCGGCCATTAGTGCTG	1143	Forward primer to construct pGL3-basic-H1/H2/H3
P4F	AAAGAGCTCTGCCGCCCTGTATGTTTCT	920	Forward primer to construct pGL3-basic-Del
P3R	AAAAAGCTTGCAACCCGCGAACGAG		Reverse primer to construct pGL3 reporters

Restriction enzyme recognition sites were in bold. *Sac*I:GAGCTC; *Hind*III:AAGCTT

were identified by visual inspection of both chromatograms with BioEdit [12]. The genotyping analyses of the four SNPs were conducted using all the 318 pigs in the study.

### Haplotype construction and statistical analysis

Pairwise tests for linkage disequilibrium (LD) and haplotype construction were performed using the SHEsis software platform (<http://analysis2.bio-x.cn/myAnalysis.php>) [13]. Normalized LD coefficient ( $D'$  value) and correlation coefficient between a pair of the alleles ( $r^2$  value) were estimated. Association of diplotypes with economic traits were tested using the MIXED procedure of SAS (SAS Institute, Cary, NC), with a model including genotype as fixed effect, sire and litter as random effects, and off-test date or weight as a covariate. Significant level was set to be 0.05.

### Plasmid construction

To construct the plasmids containing promoter region of porcine *PNPLA2* gene, the 5' end of primers were modified (highlighted in bold) to contain *Sac*I and *Hind*III restriction sites in the forward primer (P3F, P4F) and reverse primer (P3R), respectively (Table 1). The 1143 bp length of promoter region (−848 to +295 nt), which harbouring the different haplotypes (H1, H2, H3), was amplified by primer pairs P3F/P3R, and 920 bp length of promoter region (−615 to +295 nt), which lacking the haplotypes (Del), was amplified by primer pairs P4F/P3R. PCR was performed under the same condition described above. The amplified fragments were purified and quantified, and then double-digested with *Sac*I and *Hind*III and ligated into the pGL3-basic vector (Promega). All the recombinant plasmids were identified by DNA sequencing, and designated as pGL3-basic-H1, pGL3-basic-H2, pGL3-basic-H3 and pGL3-basic-Del, respectively (Fig. 1). A pGL3-promoter vector (Promega) driven by SV40 promoter served as a

positive control, and a promoter-less pGL3-basic vector as a background control.

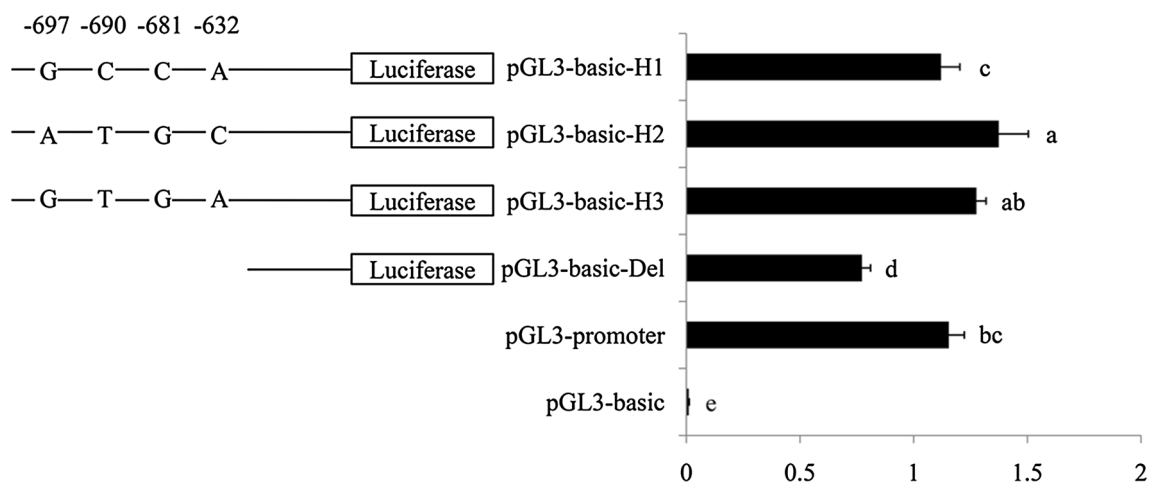
### Transfection and dual-luciferase reporter assay

Human embryonic kidney (HEK) 293 cells were cultured in Dulbecco's modified Eagle's medium (HyClone) with 10 % fetal bovine serum (Biowest) and 1 % penicillin streptomycin (Invitrogen) under 5 % CO<sub>2</sub> at 37 °C. The recombinant plasmids containing different DNA sequences of porcine *PNPLA2* promoter were transfected into HEK293 cells using lipofectamine 2000 (Invitrogen). There were six groups in the experiment: positive control group (transfected by pGL3-promoter plasmid), negative control (transfected by pGL3-basic plasmid), control group (transfected by pGL3-basic-Del plasmid) and experimental groups (transfected by pGL3-basic-H1, pGL3-basic-H2 and pGL3-basic-H3 plasmid respectively). The plasmid pRL-TK (Promega) was co-transfected into all cells for normalization. Luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) and GloMax-Multi Detection System (Promega). All experiments were performed in three independent replication and each reaction was carried out in triplicate. One-way ANOVA analysis was conducted to compare the differences among different groups.

## Results

### Identification and in silico analysis of four SNPs

A length of 3107 bp upstream sequence of porcine *PNPLA2* gene was obtained with primer pairs P1F/P1R and deposited in GenBank (JN615017). Four new SNPs were identified in this region, including A>G at −697 nt (SNP1), T>C at −690 nt (SNP2), G>C at −681 nt (SNP3) and C>A at −632 nt



**Fig. 1** Different haplotype constructs of porcine *PNPLA2* gene (left) and dual-luciferase reporter assays in HEK293 cell. Left: diagram of four promoter constructs showing three different haplotypes and haplotype deficiency as control. The number indicated the relative

nucleotide position of the four SNPs from the transcriptional start site (TSS; position +1). Right: luciferase activity was presented as the value relative to that of pRL-TK. Bars with different lowercase letters (a, b, c, d and e) indicate significantly promoter activities ( $P < 0.05$ )

(SNP4), relative to the predicted transcription start site (+1). These were all submitted to the dbSNP (NCBI) and received SNP IDs as follows: ss#1985401161: NC\_010444.3:g.161969 T>C (SNP1), ss#1985401160: NC\_010444.3:g.161962 A>G (SNP2), ss#1985401159: NC\_010444.3:g.161953 C>G (SNP3) and ss#1985401158: NC\_010444.3:g.161904 G>T (SNP4).

In-silico analysis of the SNP loci using transcription factor binding site prediction tool TESS by searching TRANSFAC Matrices is presented in Table 2. Both C allele in SNP1 and T allele in SNP4 introduced a binding site for glucocorticoid receptor (GR), while G allele in SNP4 introduced a binding site for specificity protein 1 (Sp1). An allele in SNP2 introduced binding sites for CAC-binding protein and activating protein 2 (AP-2), while G allele introduced the binding site for estrogen receptor alpha (ER $\alpha$ ). No different binding sites predicted for the different allele in SNP3.

### Genotyping of four SNPs in five pig populations

By direct sequencing of a 608 bp fragment amplified with primer pairs P2F/P2R (Table 1), the four SNPs were

detected in 318 animals of five different pig breeds (Table 3). The allele distribution revealed SNP1 and SNP4 were occurred in a chain, while SNP2 and SNP3 were chain mutations. All breeds were polymorphic at the SNP2 (SNP3) locus except Landrace breed, and all four SNPs were in agreement with Hardy–Weinberg equilibrium in Duroc breed.

### Linkage disequilibrium and haplotype construction

The LD between the four SNPs in all pig populations were estimated, which indicated that SNP1 and SNP4, SNP2 and SNP3 were in complete LD ( $D' = 1$  and  $r^2 = 1$ ) in all pig breeds (Table 4).

Haplotype-based analysis can provide higher power, precision and quality to assess the complex relationship between genetic variation and phenotypes compared to single SNP analysis [14–17]. Through estimating SNP haplotype frequencies, three haplotypes were identified, including H1(CGGT), H2(TACG) and H3(CACT) (Table 5). In the five pig populations, three haplotypes had a moderate distribution in Yorkshire breed, with frequencies of 17.6 % for H1, 38.2 % for H2 and 44.2 % for H3.

**Table 2** Prediction of transcription factors binding with porcine *PNPLA2* gene

SNP <sup>a</sup>	Position <sup>b</sup>	Allele	Putative TF	Allele	Putative TF
SNP1:g.161969 T>C	−697	T	/	C	GR
SNP2:g.161962 A>G	−690	A	CAC-binding pro; AP2	G	ER $\alpha$
SNP3:g.161953 C>G	−681	C	/	G	/
SNP4:g.161904 G>T	−632	G	Sp1	T	GR

<sup>a</sup> SNP submitted to dbSNP with positions relative to NC\_010444.3; <sup>b</sup>SNP position relative to the predicted transcription start site of porcine *PNPLA2* gene

TF transcription factor

**Table 3** Allele frequencies of four SNPs in porcine *PNPLA2* gene in five populations

Breeds	Number of pigs	SNP1(SNP4)			SNP2(SNP3)		
		T(G)	C(T)	$\chi^2$	G(G)	A(C)	$\chi^2$
Duroc	161	0.24	0.76	1.984	0.75	0.25	3.583
Landrace	63	0.78	0.22	0.007	0	1	/
Yorkshire	51	0.38	0.62	6.980**	0.18	0.82	0.321
SXHZ	19	0	1	/	0.26	0.74	3.971*
BM	24	0	1	/	0.94	0.06	9.967**

SNP1: g.161969 T&gt;C; SNP2: g.161962 A&gt;G; SNP3: g.161953 C&gt;G; SNP4: g.161904 G&gt;T

 $\chi^2_{0.05}(1) = 3.84$ ,  $\chi^2_{0.01}(1) = 6.63$ \*,\*\* Indicates significant differences at  $P < 0.05$  and  $P < 0.01$  level, respectively**Table 4** Estimated values of linkage equilibrium analysis for four SNPs in porcine *PNPLA2* gene in studied populations

SNPs	Duroc		Landrace		Yorkshire		SXS		BM	
	$D'$	$r^2$	$D'$	$r^2$	$D'$	$r^2$	$D'$	$r^2$	$D'$	$r^2$
1/2	1.000	0.952	0	0	0.999	0.132	0	0	0	0
1/3	1.000	0.952	0	0	0.999	0.132	0	0	0	0
1/4	1.000	1.000	1.000	1.000	1.000	1.000	0	0	0	0
2/3	1.000	1.000	0	0	1.000	1.000	1.000	1.000	1.000	1.000
2/4	1.000	0.952	0	0	0.999	0.132	0	0	0	0
3/4	1.000	0.952	0	0	0.999	0.132	0	0	0	0
Mean	1.000	0.968	0.167	0.167	0.999	0.421	0.167	0.167	0.167	0.167

Duroc,  $n = 161$ ; Landrace,  $n = 63$ ; Yorkshire,  $n = 51$ ; SXS Shengxian Spot pig,  $n = 19$ ; BM Bamei pig,  $n = 24$ 

SNP1: g.161969 T&gt;C; SNP2: g.161962 A&gt;G; SNP3: g.161953 C&gt;G; SNP4: g.161904 G&gt;T

However, only two haplotypes had allele frequencies greater than 5 % in the other four populations. H1, H2 and H3 were the most abundant haplotypes in Duroc (0.75), Landrace (0.78) and Chinese indigenous breeds (>0.73), respectively. In the Duroc population analysed, three haplotypes and five diplotypes were identified, and H1 and H2

accounted for 99 % of the observations. The most prevalent haplotype, H1, accounted for 74.5 % of all haplotypes, and the most common diplotype (58.4 % of samples) comprised H1 homozygotes (Table 5). In Landrace population, H2 and H3 were identified with frequencies of 77.8 and 22.2 % respectively, and the most common diplotype

**Table 5** Diplotype and haplotype frequencies of the four SNPs in 5' flanking region of porcine *PNPLA2* gene in different populations

Items	Diplotype	Observation (frequency)				
		Duroc	L	Y	SXS	BM
H1H1	CC–GG–GG–TT	94 (0.584)	0	1 (0.020)	3 (0.158)	1 (0.042)
H2H2	TT–AA–CC–GG	13 (0.081)	38 (0.603)	3 (0.059)	0	0
H3H3	CC–AA–CC–TT	1 (0.006)	3 (0.048)	7 (0.137)	12 (0.632)	22 (0.916)
H1H2	CT–GA–GC–TG	52 (0.323)	0	11 (0.216)	0	0
H1H3	CC–GA–GC–TT	0	0	7 (0.137)	4 (0.210)	1 (0.042)
H2H3	TC–AA–CC–GT	1 (0.006)	22 (0.349)	22 (0.431)	0	0

Items	Haplotype	Frequency				
		Duroc	L	Y	SXS	BM
H1	CGGT	0.745	0	0.176	0.263	0.100
H2	TACG	0.245	0.778	0.382	0	0
H3	CACT	0.010	0.222	0.442	0.737	0.900

was H2H2 (60.3 % of all investigated samples). Chinese indigenous breeds SXHZ and BM had similar haplotype distributions, which comprising a high frequent haplotype H3 and a low frequent haplotype H1, and the most common diplotype was H3H3.

### Haplotype association analysis

Associations of the three diplotypes (H1, H2 and H3) with fat deposition related traits were analyzed in three pig populations as shown in Table 6, in which rare diplotypes (frequency < 0.03) were excluded from further analyses.

In Duroc breed, H1 and H2 were the main haplotypes. There were significant associations between the diplotypes and FCR and BF1 ( $P < 0.05$ ), and Duroc individuals with H1H1 diplotype exhibited the lowest FCR ( $P < 0.05$ ), while H2H2 had the thickest BF1 ( $P < 0.05$ ). Duroc pigs with H1H1 diplotype had a lower FCR ( $-0.19\%$ ,  $P < 0.05$ ) but a thicker BF1 ( $+0.23$  mm,  $P > 0.05$ ) than pigs with H1H2 diplotype, and a lower FCR ( $-0.24\%$ ,  $P > 0.05$ ) and a thinner BF1 ( $-2.26$  mm,  $P < 0.05$ ) than pigs with H2H2 diplotype. Pigs with H1H2 diplotype had a similar FCR ( $-0.05\%$ ,  $P > 0.05$ ) but a significant decrease in BF1 ( $-2.49$  mm,  $P < 0.05$ ) compared with pigs with H2H2 diplotype. There were no significant differences among different diplotypes for the other traits

including TDG, BF2, MD2 and LMP. H3 was present at too low a frequency (0.01) in Duroc breed population for subsequent analysis.

In Landrace breed analyzed, only two haplotypes (H2, H3) and three diplotypes were identified. There were significant associations between the diplotypes and BF1 and MD ( $P < 0.05$ ), extremely significant with LMP ( $P < 0.05$ ). Landrace individuals with H2H3 had lower BF1 ( $P < 0.05$ ), higher MD ( $P < 0.05$ ) and LMP ( $P < 0.05$ ) than those with diplotype H2H2 and H3H3. No significant differences were observed among diplotypes for the other traits including TDG, FCR and BF2.

In Yorkshire breed analyzed, three haplotypes were identified and five diplotypes were selected for further association because of the rare frequency of H1H1. However, there were no significant differences among different diplotypes for all the traits measured.

### Functional analysis of different haplotypes

To investigate the potential effects of different haplotypes on porcine *PNPLA2* transcription, the 1143 bp fragment of 5'-flanking region from  $-848$  to  $+295$  nt was cloned into a luciferase reporter vector. Four constructs differing only at sites  $-697$  (SNP1),  $-690$  (SNP2),  $-681$  (SNP3) and  $-632$  (SNP4) were obtained, three of which harbouring the

**Table 6** Effects of the different diplotypes in porcine *PNPLA2* gene on economic traits

Diplotype	N	TDG (kg)	FCR	BF1 (mm)	BF2 (mm)	MD (mm)	LMP (%)
Duroc							
H1H1	94	0.76 ± 0.01	2.43 ± 0.04 <sup>a</sup>	17.31 ± 0.31 <sup>a</sup>	15.89 ± 0.35	46.67 ± 1.00	60.74 ± 0.33
H1H2	52	0.73 ± 0.01	2.62 ± 0.06 <sup>b</sup>	17.08 ± 0.52 <sup>a</sup>	16.38 ± 0.60	48.93 ± 1.70	60.35 ± 0.56
H2H2	13	0.71 ± 0.02	2.67 ± 0.13 <sup>ab</sup>	19.57 ± 0.81 <sup>b</sup>	16.59 ± 0.93	44.05 ± 2.61	59.26 ± 0.86
<i>P</i>		0.081	0.0302	0.027	0.6885	0.2553	0.2966
Landrace							
H2H2	38	0.70 ± 0.01	2.35 ± 0.04	14.34 ± 0.47 <sup>a</sup>	13.33 ± 0.46	45.17 ± 0.97 <sup>a</sup>	63.10 ± 0.35 <sup>A</sup>
H2H3	22	0.72 ± 0.02	2.46 ± 0.05	12.74 ± 0.51 <sup>b</sup>	11.81 ± 0.49	48.90 ± 1.04 <sup>b</sup>	64.93 ± 0.37 <sup>B</sup>
H3H3	3	0.73 ± 0.03	2.26 ± 0.08	15.40 ± 0.81 <sup>a</sup>	12.62 ± 0.78	46.15 ± 1.66 <sup>ab</sup>	62.86 ± 0.60 <sup>A</sup>
<i>P</i>		0.5372	0.1070	0.0184	0.0942	0.0419	0.0023
Yorkshire							
H1H2	11	0.66 ± 0.01	2.60 ± 0.09	15.00 ± 1.10	16.16 ± 1.16	49.38 ± 3.24	62.52 ± 0.88
H1H3	7	0.69 ± 0.01	2.65 ± 0.10	14.76 ± 1.06	13.76 ± 1.12	52.52 ± 3.12	63.77 ± 0.85
H2H2	3	0.73 ± 0.03	2.86 ± 0.25	/	/	/	/
H2H3	22	0.67 ± 0.01	2.55 ± 0.05	14.51 ± 0.70	14.10 ± 0.73	54.09 ± 2.06	64.03 ± 0.56
H3H3	7	0.67 ± 0.02	2.77 ± 0.18	17.78 ± 1.50	16.83 ± 1.58	55.52 ± 4.42	61.65 ± 1.21
<i>P</i>		0.0508	0.3534	0.4487	0.3774	0.2457	0.3037

Different letters denoting significant difference between groups: a, b, c, d, e,  $P < 0.05$ . All data indicated as least square mean values ( $\pm$ SE) TDG average daily gain between 30 and 100 kg body weight, FCR feed conversion ratio, BF1 backfat thickness between 3rd and 4th last lumbar vertebrae, 7 cm from the midline, BF2 and MD backfat thickness and MD between 3rd and 4th last rib and 7 cm from the midline, LMP estimated lean meat percentage

identified haplotypes, and the remaining control was short of the sequence from  $-848$  to  $-615$  nt. As shown in Fig. 1, functional analysis on *PNPLA2* promoter haplotypes indicated that the DNA fragment of *PNPLA2* promoter ( $-848$  to  $+295$  nt) harbouring the four SNPs had a strong activity of promoter, compared with positive control transected with pGL3-promoter. The activity of *PNPLA2* promoter was influenced by the polymorphisms of the four SNPs. The mean relative luciferase activity (RLA) of H2 (TACG) and H3 (CACT) haplotype were significantly higher than that of H1 (CGGT) haplotype ( $P < 0.05$ ), but there was no significant difference with the RLA between H2 and H3 haplotype ( $P > 0.05$ ). The RLA of pGL3-basic-H2 was the highest among the all vectors, even significantly higher than pGL3-promoter containing SV40 promoter and enhancer, which indicated H2 might cause a strong positive regulatory element. As an experimental control, the activity of pGL3-basic-Del (containing the sequence from  $-615$  to  $+295$ nt but lacking the variation region) was significantly higher than pGL3-basic ( $P < 0.05$ ), which indicated the DNA fragment of *PNPLA2* promoter (from  $-615$  to  $+295$ nt) also had a high activity of promoter. However, the activity of pGL3-basic-Del was lower than the other three recombinant vectors containing the variation region from  $-848$  to  $-615$ nt, which indicated the DNA fragment (from  $-848$  to  $-615$  nt) had an activity of enhancer.

## Discussion

*PNPLA2* specifically catalyzes the first step in triglyceride hydrolysis in mammalian cells. Therefore, *PNPLA2* gene has been an attractive candidate gene for fat deposition related traits in livestock species. Mutations in regions upstream of transcription initiation site may lead to variation in the level or timing of gene expression by destroying or increasing the number of transcription factor binding sites [11, 18, 19]. In this study, we first reported and confirmed four SNPs in the promoter region of *PNPLA2* gene in 318 pigs by direct PCR amplification and DNA sequencing methods, revealed significant genetic effects between the mutations and growth and fat deposition traits of pigs, and identified the transcriptional activities for different haplotypes.

Comparison analysis revealed that there are four contiguous SNPs on the upstream of transcription initiation site of porcine *PNPLA2* gene. It is considered that a haplotype analysis can be more practical for most complex traits, because a single SNP usually function in coordination with other SNPs to manifest an effect [20, 21]. In the study, haplotypes were constructed using four SNPs of porcine *PNPLA2* gene, among which SNP1 and SNP4 are linked completely, so as for SNP2 and SNP3. H1 (CGGT)

and H2 (TACG) occupy the whole Duroc population detected and carry the opposite alleles at all the four loci. H2 (TACG) and H3 (CACT) carry the same alleles at SNP2 and SNP3, but opposite at SNP1 and SNP4.

Statistical analysis is one of the ways to clarify the genetic effects. Significant associations between *PNPLA2* haplotypes and backfat thickness were observed in Duroc and Landrace. Haplotype analysis showed in Duroc, individuals with H1 haplotype had the lower FCR and BF1 ( $P < 0.05$ ), which indicated H1 was the preferential haplotype for FCR and backfat thickness traits in Duroc. Interestingly, the decrease effect of H1 on backfat thickness is reflected by the two predicted binding sites for GR and ER $\alpha$  introduced in SNP1 and SNP4. However, individuals with H1 were absent in Landrace, and it showed the preferential heterozygote of H2H3 for backfat thickness (BF2), muscle depth (MD) and LMP in this population. The genetic effects of these haplotypes were not consistent across breed studied. Although the genetic effects should be verified in a larger scale of pig populations, the associations between *PNPLA2* promoter haplotypes and phenotypes suggest the possibility of differential promoter activity.

In vitro study, three haplotypes, which consists of four nucleotide variants from  $-848$  to  $-615$  nt in porcine *PNPLA2* promoter, showed different basal transcriptional activities in HEK293 cells. It indicated that the haplotype H2 (TACG), which was the most frequent in Landrace, had the highest transcriptional activity, followed by H3(CACT) which mainly existed in Yorkshire, SXH and BM pigs, and then H1 (CGGT) which mainly existed in Duroc. This suggests that the different haplotypes caused by nucleotide variants could change binding affinity to transcription factors individually or interactively. From in vitro data, a lower transcriptional activity of H1 (compared to H3) could be due to differences in ER $\alpha$  binding site at G allele in SNP2 and CAC-binding protein and AP2 binding sites at A allele in SNP2, while a higher transcriptional activity of H2 (compared to H3) could due to differences in GR binding site at G allele in SNP1 (T allele in SNP4) and Sp1 binding sites at G allele in SNP4.

In H1 haplotype, two GR binding sites and an ER $\alpha$  binding site were predicted to be introduced. It has been demonstrated that *PNPLA2* gene expression is induced by glucocorticoids (GCs), one of the hormones involved in the response of an organism to fasting [3]. Dexamethasone (GC) can directly stimulate lipolysis by up-regulating mRNA and protein levels of *PNPLA2* in a dose and time responsive manner [22]. Therefore, in present study, the action of GCs on HEK293 cells may be mediated by GR, which is a widely expressed nuclear receptor in virtually all tissues. The GR homodimer binds to the glucocorticoid response elements (GREs) and then stimulates *PNPLA2*

expression [23, 24]. Fortunately, we found GREs were predicted to be introduced in H1 and H3 or abolished in H2 in the 5'-promoter region of porcine *PNPLA2* gene due to the mutations. In vivo, GC can be induced in the stress condition such as being fasted, and then pigs with C allele both in SNP1 and SNP4 in *PNPLA2* gene could be more easily induced by GC and GR, as a result, the enhanced lipolysis activities of *PNPLA2* would gradually result in their thinner backfat thicknesses than pigs with T allele both in SNP1 and SNP4.

Furthermore, evidence has demonstrated a pivotal role of estrogens and estrogen receptors (ERs) in fat formation and metabolism [25]. Estrogen (E2) and ER $\alpha$  have been shown to inhibit adipogenesis [26], and the role of estrogen in adiposity is mediated mainly through ER $\alpha$  [27]. It has been demonstrated that E2 signals via ER $\alpha$  to regulate *PNPLA2* and perilipin-mediated lipid metabolism and droplet size in adipocytes [28]. In our study, the ER $\alpha$  binding site introduced in H1 might be recognized by estrogenic complex and lead to the transcriptional activation of *PNPLA2* in estrogen stimulation in vivo. Therefore, the GC-GR and E2-ER $\alpha$  actions in vivo might explain the reason why haplotype H1 had thinner backfat thickness than Duroc pigs with H2. Studies of tissue-specific differences in GC and estrogen metabolism and their receptor-mediated actions will be helpful to understand the effects of introduced GR and ER $\alpha$  binding sites on porcine *PNPLA2* regulation.

However, while in vitro, HEK293 cells in the present study were not being fasted or treated with GC or estrogen. So there were no promotion effects of GR binding sites or ER $\alpha$  binding site on *PNPLA2* transcription. This might explain the reason why haplotype H1 had a relatively low transcriptional activity in HEK293 cells. In vitro results might not represent for in vivo results under the environment with various hormones such as GC and E2.

There were other transcriptional factor binding sites predicted in H2 and H3, such as Sp1, AP-2 and CAC-binding proteins, which are broadly expressed transcription factors and capable of contributing to basal promoter activity. GR binding sites in H1 and H3 might not be induced by GC in HEK293 cells in vitro. Therefore, we considered it may explain the reason why H2 had the highest transcriptional activity in HEK293 cells.

Sp1, as a basal transcription factor, is a zinc finger transcription factor with important roles in inducing the expression of genes associated with lipid metabolism [29–32]. It can activate the transcription of many cellular genes that contain putative CG-rich Sp-binding sites in their promoters [33].

AP-2 is an inducible transcription factor and represents a family of five sequence-specific DNA-binding proteins [34]. It can regulate genes containing AP-2 sites in their

promoters and play roles in various important biological functions, such as apoptosis and cell growth [35, 36].

CAC box in the promoter could be recognized by CAC-binding proteins encoded by genes with 3'UTR containing motif GRGGSTGGG [37]. It was reported that deletion of a CAC box could result in a drastic decrease in transcription of the porphobilinogen deaminase (*PBGD*) promoter [37].

Moreover, in the upstream region of porcine *PNPLA2* gene, a number of transcription factor binding sites were predicted such as CREBP1, MyoD, C/EBP, nuclear factor 1 (NF-1). These transcription factors might also be involved in the functioning of the *PNPLA2* promoter besides the five transcription factors above.

Further biochemical and molecular characterization using electrophoretic mobility shift assays and DNase I protection assays could elucidate key elements, as well as further research on the transcriptional level of *PNPLA2* gene in pigs may provide useful information on lipometabolism and suggestions on marker-assisted selection as well.

## Conclusion

In conclusion, four functional polymorphisms were identified in the promoter region of porcine *PNPLA2* gene. The allele order of C–G–G–T, corresponding to SNP1–4 or H1, had consistent trends in decreasing FCR and backfat thickness in Duroc pigs, but had a lower basal transcriptional activity in HEK293 cells in vitro. These results suggest that the four SNPs were potential DNA markers for backfat thickness in Duroc pig breeding program and had a regulatory role in *PNPLA2* transcription. The in vivo and in vitro experiments should be in more way in further study.

**Acknowledgments** The work was supported by National Natural Science Foundation of China (31201783), National Major Project for Production of Transgenic Breeding Grant (2014ZX0800945B), and National Natural Science Foundation of Zhejiang Province (LY15C170003, Y13C170012).

## References

- Zimmermann R, Strauss JG, Haemmerle G, Schoiswohl G, Birner-Gruenberger R, Riederer M, Lass A, Neuberger G, Eisenhaber F, Hermetter A, Zechner R (2004) Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase. *Science* 306:1383–1386
- Jenkins CM, Mancuso DJ, Yan W, Sims HF, Gibson B, Gross RW (2004) Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. *J Biol Chem* 279:48968–48975



3. Villena JA, Roy S, Sarkadi-Nagy E, Kim KH, Sul HS (2004) Desnutrin, an adipocyte gene encoding a novel patatin domain-containing protein, is induced by fasting and glucocorticoids: ectopic expression of desnutrin increases triglyceride hydrolysis. *J Biol Chem* 279:47066–47075
4. Kershaw EE, Hamm JK, Verhagen LA, Peroni O, Katic M, Flier JS (2006) Adipose triglyceride lipase function, regulation by insulin, and comparison with adiponutrin. *Diabetes* 55:148–157
5. Smirnova E, Goldberg EB, Makarova KS, Lin L, Brown WJ, Jackson CL (2006) ATGL has a key role in lipid droplet/adiposome degradation in mammalian cells. *EMBO Rep* 7(1):106–113
6. Chen JF, Dai LH, Xu NY, Xiong YZ, Jiang SW (2006) Assignment of the patatin-like phospholipase domain containing 2 gene (*PNPLA2*) to porcine chromosome 2p17 with radiation hybrids. *Cytogenet Genome Res* 112(3–4):342G
7. Rattink AP, Faivre M, Jungerius BJ, Groenen MA, Harlizius B (2001) A high-resolution comparative RH map of porcine chromosome (SSC) 2. *Mamm Genome* 12:366–370
8. Schoenborn V, Heid IM, Vollmert C, Lingenhel A, Adams TD, Hopkins PN, Illig T, Zimmermann R, Zechner R, Hunt SC, Kronenberg F (2006) The ATGL gene is associated with free fatty acids, triglycerides, and type 2 diabetes. *Diabetes* 55(5):1270–1275
9. Fischer J, Negre-Salvayre A, Salvayre R (2007) Neutral lipid storage diseases and ATGL (adipose triglyceride lipase) and CGI-58/ABHD5 (alpha-beta hydrolase domain-containing 5) deficiency: myopathy, ichthyosis, but no obesity. *Med Sci (Paris)* 23(6–7):575–578
10. Dai LH, Xiong YZ, Chen JF, Deng CY, Zuo B, Xu DQ, Lei MG, Zheng R, Li FE, Jiang SW (2011) Molecular characterization and association analysis of porcine adipose triglyceride lipase (*PNPLA2*) gene. *Mol Biol Rep* 38:921–927
11. Chorley BN, Wang X, Campbell MR, Pittman GS, Nouredine MA, Bell DA (2008) Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res Mutat Res* 659(1–2):147–157
12. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
13. Shi YY, He L (2005) SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 15:97–98
14. Tan Q, Christiansen L, Christensen K, Bathum L, Li S, Zhao JH, Kruse TA (2005) Haplotype association analysis of human disease traits using genotype data of unrelated individuals. *Genet Res* 86(3):223–231
15. Li J, Zhou Y, Elston RC (2006) Haplotype-based quantitative trait mapping using a clustering algorithm. *BMC Bioinform* 7:258
16. Rieder MJ, Taylor SL, Clark AG, Nickerson DA (1999) Sequence variation in the human angiotensin converting enzyme. *Nat Genet* 22:59–62
17. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
18. Wang YH, Byrne KA, Reverter A, Harper GS, Taniguchi M, McWilliam SM, Mannen H, Oyama K, Lehnert SA (2005) Transcriptional profiling of skeletal muscle tissue from two breeds of cattle. *Mamm Genome* 16(3):201–210
19. Wang X, Tomso DJ, Chorley BN, Cho HY, Cheung VG, Kleeburger SR, Bell DA (2007) Identification of polymorphic antioxidant response elements in the human genome. *Hum Mol Genet* 16(10):1188–1200
20. Singh M, Singh P, Juneja PK, Singh S, Kaur T (2011) SNP–SNP interactions within APOE gene influence plasma lipids in post menopausal osteoporosis. *Rheumatol Int* 31:421–423
21. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. *Nat Genet* 29:229–232
22. Xu C, He JH, Jiang HF, Zu LX, Zhai WJ, Pu SS (2009) Direct effect of glucocorticoids on lipolysis in adipocytes. *Mol Endocrinol* 23:1161–1170
23. Oakley RH, Cidlowski JA (2011) Cellular processing of the glucocorticoid receptor gene and protein: new mechanisms for generating tissue-specific actions of glucocorticoids. *J Biol Chem* 286:3177–3184
24. Lee MJ, Pramyothin P, Karastergiou K, Fried SK (2014) Deconstructing the roles of glucocorticoids in adipose tissue biology and the development of central obesity. *Biochim Biophys Acta* 1842(3):473–481
25. Xiao J, Wang NL, Sun B, Cai GP (2010) Estrogen receptor mediates the effects of pseudoprotodiocsin on adipogenesis in 3T3-L1 cells. *Am J Physiol* 299(1):C128–C138
26. D'Eon TM, Souza SC, Aronovitz M, Obin MS, Fried SK, Greenberg AS (2005) Estrogen regulation of adiposity and fuel partitioning. Evidence of genomic and non-genomic regulation of lipogenic and oxidative pathways. *J Biol Chem* 80:35983–35991
27. Metz L, Gerbaix M, Masgrau A, Guillet C, Walrand S, Boisseau N, Boirie Y, Courteix D (2016) Nutritional and exercise interventions variably affect estrogen receptor expression in the adipose tissue of male rats. *Nutr Res* 36(3):280–289
28. Wend K, Wend P, Drew BG, Henvener AL, Miranda-Carboni GA, Krum SA (2013) ER $\alpha$  regulates lipid metabolism in bone through ATGL and perilipin. *J Cell Biochem* 114(6):1306–1314
29. Solomon SS, Majumdar G, Martinez-Hernandez A, Raghov R (2008) A critical role of Sp1 transcription factor in regulating gene expression in response to insulin and other hormones. *Life Sci* 83(9–10):305–312
30. Huang C, Xie K (2012) Crosstalk of Sp1 and Stat3 signaling in pancreatic cancer pathogenesis. *Cytokine Growth Factor Rev* 23(1–2):25–35
31. Zhu J, Sun Y, Luo J, Wu M, Li J, Cao Y (2015) Specificity protein 1 regulates gene expression related to fatty acid metabolism in goat mammary epithelial cells. *Int J Mol Sci* 16(1):1806–1820
32. Zhu JJ, Luo J, Xu HF, Wang H, Looor JJ (2016) Short communication: altered expression of specificity protein 1 impairs milk fat synthesis in goat mammary epithelial cells. *J Dairy Sci* 99(6):4893–4898
33. Vizcaíno C, Mansilla S, Portugal J (2015) Sp1 transcription factor: a long-standing target in cancer chemotherapy. *Pharmacol Ther* 152:111–124
34. Eckert D, Buhl S, Weber S, Jäger R, Schorle H (2005) The AP-2 family of transcription factors. *Genome Biol* 6(13):246
35. Zeng YX, Somasundaram K, El-Deiry WS (1997) AP2 inhibits cancer cell growth and activates p21WAF1/CIP1 expression. *Nat Genet* 15:78–82
36. Hilger-Eversheim K, Moser M, Schorle H, Buettner R (2000) Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. *Gene* 260(1–2):1–12
37. Mignotte V, Wall L, Deboer E, Grosveld F, Romeo PH (1989) Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res* 17(1):37–54