SNPs identification in swine leptin 5′ flanking region and transcriptional activity of naturally occurring promoter haplotypes

Alessandra Crisà,1 Mariasilvia D’Andrea,2 Daniela Willems,3 Fabio Pilla,2 Alessio Valentini1

1Centro di ricerca per la produzione delle carni e il miglioramento genetico, Monterotondo (RM), Italy
2Dipartimento di Scienze Animali Vegetali e dell’Ambiente, Università del Molise, Campobasso, Italy
3Dipartimento per la innovazione nei sistemi biologici, agroalimentari e forestali, Università della Tuscia, Viterbo, Italy

Abstract

Leptin is an adipocyte-derived hormone which acts as a major regulator for food intake and energy homeostasis. Previous studies have focused on the association between polymorphisms in the coding region of the leptin (LEP) pig gene and economically important traits. The present work aimed to study the leptin gene at the promoter region to search for polymorphisms that could lead to different gene expression regulation. In the 1213 bp 5′ gene flanking region, we identified 17 new polymorphisms that could be partly due to differences in the genetic background of the breeds analyzed and to the relatively scarce attention paid to the regulatory regions of the gene.

Although previous studies have focused on associations between polymorphisms in the coding regions of the leptin gene and economically important traits, studies in humans and other species have shown that polymorphisms in the leptin promoter may be of major importance. In cattle, single marker association analyses have been carried out to evaluate the relationship between different genotypes of two SNPs in the leptin promoter region and serum leptin concentration, growth rate and backfat thickness. In a work by Stachowiak et al. (2007), four polymorphisms were found in a 245 bp fragment of the pig promoter region but the authors did not find evidence for an association of the LEP promoter genotype with the fatness traits analyzed.

The aim of this work was to study a region upstream of the 5′ of the transcriptional start site which is longer than those previously investigated, to search for SNPs that could be associated with different leptin gene transcriptional activity. The analysis of the 5′ flanking region of the gene showed the presence of many polymorphisms both in Casertana (CT) and Large White (LW) pig breeds. These polymorphisms were arranged in phased alleles (haplotype), two of which were examined in transient transfection reporter gene assays.

Experimental population

A total of 30 animals were analyzed: 15 Casertana (CT) and 15 Large White (LW) pigs of both sexes. The Casertana pig is a local breed from Southern Italy characterized by slow growth and massive accumulation of backfat. In the past, the breeding of this species has not been encouraged mainly because of economic viability. This is due to the long time needed for them to reach their mature body weight and also to consumer preference. Large White is a cosmopolitan breed with fast growth which is bred for lean meat.

Casertana pigs were chosen from three farms at an experimental farm belonging to the Parco Scientifico e Tecnologico del Molise. The farm was established by collecting founder animals from all the areas where Casertana were reared. Since the effective number of the population was very low, it was impossible to avoid kinship within the selected animal group. Large White were purchased from commercial farms avoiding strictly related individuals. All animals were raised outdoors in the same environmental conditions and fed the same diet. The animals were slaughtered at 11 months old. On the day of slaughter, they were transported to a commercial slaughterhouse where they were kept for a minimum of 12 h
prior to slaughtering. The liver tissue was collected at slaughter and stored at -20°C in RNA later. All animals were reared and slaughtered according to European directives and laws.

DNA extraction and PCR amplification

Genomic DNA was isolated from the liver of the 30 animals studied using the Wizard Genomic kit (Promega, Madison, WI, USA). PCR primers were designed on the basis of the available published pig leptin sequence AF492499 by using DNAMAN software (Version 4.15, Lynnon BioSoft, Inc., Quebec, Canada). The primer sequences, amplicon sizes and positions are reported in Table 1.

PCRs for the 596 and 637 bp amplicons were performed in a 50 µl reaction containing 100 ng of genomic DNA, 40 pmol of each primer, 100 µM each of the four dNTPs, 10 µl 5X GoTaq reaction buffer and 2.5 U GoTaq DNA polymerase (Promega, Madison, USA). Amplification touchdown conditions were: 95°C for 5 min followed by 14 cycles at 94°C for 30 sec, then at 67°C for 30 sec with a -0.5°C decrement per cycle and at 72°C for 1 min; the subsequent 19 cycles were at 94°C for 30 sec followed by 60°C for 30 sec and 72°C for 1 min. A final step at 72°C for 10 min was added. Amplicons were first recovered from the gel using the Wizard SV Gel and PCR Purification Kit (Promega, Madison, USA) and subsequently cloned into the pCR-Blunt vector (Invitrogen, Paisley, UK), quantified using the Quant-iT™ Picogreen dsDNA binding Reagent and Kits (Molecular Probes, Eugene, USA) and DNA was bidirectionally sequenced using the DCTS kit and the CEQ8800 sequencer from Beckman; some sequences were double-checked with an outsourcing service (MWG, Ebersberg, Germany).

Table 1. Primers used to obtain three amplicons of the pig leptin 5’-flanking region.

| Primer name | Sequences | Size (bp) | Position on AF492499 |
|-------------|-----------|-----------|---------------------|
| Lep1Fw      | 5’AGATCTTTGGAAATATGCGAAGGC3’ | 637 | 4463 - 4486 |
| Lep1Rw      | 5’GGACAGGGGTCTGCTTCTTTGC3’ | 637 | 4463 - 4486 |
| Lep2Fw      | 5’GACAAGAAGGACGCCCTTTGCC3’ | 596 | 5082 - 5101 |
| Lep2Rw      | 5’CCCTCTTATAGACGCGGAGG3’ | 596 | 5082 - 5101 |
| Lep1Rw      | 5’AGATCTTTGGAAATATGCGAAGGC3’ | 1213 | 4463 - 4486 |
| Lep2Rw      | 5’CCCTCTTATAGACGCGGAGG3’ | 1213 | 5082 - 5101 |

Constructions and cell transfections

The leptin promoter was amplified by PCR from a heterozygous CT animal and was then linked to the luciferase gene in the pGL3 basic vector for the in vitro reporter gene assay. PCR amplification of the 1213bp promoter fragment was performed in a 20 µl reaction containing 120 ng of genomic DNA, 15 pmol of each primer, 100 µM each of the four dNTPs, 0.6 µl DMSO, 4 µl 5X HF reaction buffer and 0.4 U of the proofreading DNA polymerase Phusion (Fynzyme, Vaanta, Finland). The thermal conditions were: 98°C for 3 min followed by 30 cycles at 98°C for 10 sec, at 65°C for 30 sec and at 72°C for 1 min. A final step at 72°C for 10 min was added. Amplicons were first recovered from the gel using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, USA) and subsequently mobilized into the Smal digested and dephosphorylated promoterless pGL3 basic vector containing the luciferase reporter gene (Promega). The transformed plasmids were sequenced to determine the sequence fidelity of the synthetic products and the correct orientation; the DNA for transfections was prepared as E. coli cultures using the GenElute™ HP Endotoxin-Free Plasmid Maxiprep Kit (Sigma, St. Louis, USA). Two constructs were used, namely LEPF3 and LEPF46, containing phased genotypes from both starting chromosomes were used in transfection experiments.

Chinese Hamster Ovary (CHO) cells were grown in monolayer culture at 37°C in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum. For transient transfection, cells were seeded into 24 multiwell plates at a density of 8x10^5 cells/well. After a 24 h attachment period, the cells were cotransfected, using Lipofectamine 2000 (Invitrogen, Paisley, UK), with 200 ng of promoter-Firefly luciferase plasmids (LEPF3 or LEPF46) and 50 ng pRG-TK vector containing the Renilla luciferase gene to normalize for the efficiency of transfection. Experimental controls included cells transfected with the pCMV-luc vector (positive control to test the system; data not shown) and cells transfected with the pGL3 basic vector (negative control). The cultures were incubated for a further 24 h, then rinsed with PBS and lysed in 100 µl of 1x passive lysis buffer for 10 min. Aliquots (20 µl) of whole cell extracts were analyzed sequentially for Firefly and Renilla luciferase activities using a Dual Luciferase Reporter Assay System (Promega), and the ratio of Firefly to Renilla (RLU) was taken to represent the normalized firefly luciferase activity. Each data point was repeated five times in a single experiment with the results averaged.

Statistical analysis

Statistical analysis was conducted using the software SYSTAT 10.2 (http://www.systat.com). The in vitro association between haplotype and transcription efficiency was carried out using the model

\[ y = \mu + h + \exp + \varepsilon \]

where: \( y=\text{Firefly/Renilla ratio} \), \( \mu \) is the overall mean, \( h \) is the tested haplotype, \( \exp=\text{experiment replicate} \) and \( \varepsilon \) is the residual error.

Results and discussion

Promoter sequence analysis

The analysis of the leptin promoter region in 30 animals (15 CT and 15 LW) showed a total of 17 new polymorphic variants, including 14 SNPs, a 14-base deletion, a three-base deletion and a two-base variation. All the identified mutations were submitted to the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP) and rs numbers were assigned (Table 2). In all tested animals, the SNPs were in the heterozygous state and the LW breed was the most heterozygous. The 99307493 SNP was the least frequent, present in only 3 CT animals. In our sample, we were not able to detect the SNPs reported by Stachowiak et al. (2007). The examined sequence shows a high SNP density in a relatively short sequence as also reported in dairy cattle (Lieber et al., 2005).

PCR-RFLP tests, useful for association studies in larger populations, have been developed for 6 of the 17 polymorphisms. The SNPs 99307477, 99307482 and 99307485 were detected by Tsp45I, HpyCH4V and TaqI diges-
From the genotypes of the 30 animals at the 17 SNPs positions, 5 haplotypes were inferred in silico and are reported in Table 3. Haplotype 1 had a higher frequency in CT than in LW; haplotype 2 was found only in CT while haplotypes 4 and 5 were only in the LW breed; haplotype 3 had a similar frequency in both breeds. Comparing our sequence with those from bovine (AB070368), human (U435589) and mouse (U36238) sources, we found an overall similarity of 60%, 52% and 49%, respectively. A 73% similarity was found in the alignment with the bovine sequence over the last 600 nucleotides, a region close to the transcriptional starting point (TSS). The most 3’ end of our 1213 bp amplicon was located -20bp from the putative pig TSS (Bidwell et al., 1997) compared with human (Gong et al., 1996) and -23bp compared with bovine (Taniguchi et al., 2002) and mouse (He et al., 1995).

The analysis of transcription factor binding sites relative to the 1213bp promoter sequence (i.e. the most frequent haplotype 1) detected a large number of putative recognition sites that were lost if the alternative allele was present (haplotype 3). Table 4 shows the identified sites with a matrix similarity match greater than 0.80. Analysis of the mutated sequence (haplotype 3) found binding sites not present in the haplotype 1, e.g. octamer binding protein, respectively, on the 637 bp amplicon. The SNP sites 99307489, 99307490 and 99307492 were detected by FokI, Nael and XmalIII digestion, respectively, on the 596 bp amplicon (data not shown).

Table 2. Newly identified polymorphisms in the pig leptin promoter. SNP identification and position on the AF492499 sequence, MAF in the CT and LW breeds, and assigned ss numbers at NCBI dbSNP.

| Position on AF492499 | CT   | MAF | LW   | dbSNP_ss |
|---------------------|------|-----|------|----------|
| g.4532_4545delCTTGAGACAATGGGC | 0.1  | 0.5 | 99307477 |
| g.4565C>A;g.4566T>G | 0.1  | 0.5 | 99307478 |
| g.4594A>G | 0.1  | 0.5 | 99307479 |
| g.4620C>A | 0.1  | 0.5 | 99307480 |
| g.4650_4652delITCC | 0.1  | 0.5 | 99307481 |
| g.4699T>A | 0.1  | 0.5 | 99307482 |
| g.4828A>G | 0.1  | 0.5 | 99307483 |
| g.4841T>C | 0.1  | 0.5 | 99307484 |
| g.4885T>A | 0.1  | 0.5 | 99307485 |
| g.4924G>A | 0.1  | 0.5 | 99307486 |
| g.5112T>G | 0.1  | 0.5 | 99307487 |
| g.5127G>C | 0.1  | 0.5 | 99307488 |
| g.5202C>T | 0.1  | 0.5 | 99307489 |
| g.5344A>G | 0.1  | 0.5 | 99307490 |
| g.5474C>A | 0.1  | 0.5 | 99307491 |
| g.5397T>C | 0.1  | 0.5 | 99307492 |
| g.5413C>A | 0.1  | 0 | 99307493 |

Table 3. Inferred haplotypes for the leptin promoter. For each of the five Hapl.id, we report the number of the different haplotypes inferred in the two breeds considering both chromosomes for each animal (Hapl. Freq.) and the nucleotide variations, identified with the ss number relative to the position on the AF492499 sequence (Haplotype definition).

| Hapl. id. | Haplotype definition |
|-----------|----------------------|
| CT | LW |
| 1 | 24 | 12 | / CA A C / T A T T G T G C A C T C |
| 2 | 3 | 0 | / CA A C / T A T T G T G C A C T A |
| 3 | 3 | 4 | del14 TG G A del3 A G CA A G C T G A C C |
| 4 | 0 | 7 | / CA A C / T A T T A G C T G A C C |
| 5 | 0 | 7 | del14 TG G A del3 A G CA A G C T G A C T C |

Table 4. Putative transcription factor binding sites in the 1213 bp pig leptin promoter.

| Factor | Position | Strand | Sequence<sup>o</sup> | SNP<sup>o</sup> |
|--------|----------|--------|----------------------|--------------|
| DM domain-containing transcription factors | 66 | 86 | + | tgtgcttCTGACagggctt |
| Activator/repressor binding to transcription initiation site | 67 | 87 | - | caagccCAAttcaagaccc |
| Sox/Sry-sex/testis determining and related HMG box factors | 91 | 80 | - | ttaggCCAGttcaggtgctttg |
| Estrogen response elements | 75 | 93 | - | tharrAAAGGgagtcc |
| Vertebrate steroidogenic factor | 79 | 91 | - | accaCAGggcc |
| Krueppel-like transcription factors | 98 | 114 | - | aataasgGAGggtcctt |
| AT rich interactive domain factor | 185 | 201 | - | ttagggGGGggagact |
| Abdominal B-type homeodomain transcription factors | 121 | 139 | - | aaacctgAgAattt |
| Human and murine ETS1 factors | 123 | 139 | + | ggggggggGAAacctaat |
| GC-Box factors SPU/GC | 198 | 188 | - | ggggggggGAGGaa |
| Bicoid-like homeodomain transcription factors | 183 | 197 | - | gggggggggGAGCaa |
| Ets and related proteins | 230 | 246 | + | thtgTAgActaat |
| Zinc binding protein factors | 454 | 470 | + | tggattGGGtgctt |
| Evi1-myeloid transforming protein | 652 | 674 | + | gagaacctgCPTCAGagggctt |
| AP4 and related proteins | 732 | 748 | + | ccggaggGATGaggaaa |
| Myoblast determining factors | 390 | 920 | - | agaagtCTGCGtgctt |
| Activator, mediator- and TBP-dependent core promoter element | 926 | 942 | - | gggGGCGacgccc |
| for RNA polymerase II transcription from TATA-less promoters | 933 | 951 | - | ggGTTGCGgaccc |

*Basepairs in capital letters denote the core sequence (http://www.genomatix.de/online_help/help_matinspector/matrix_help.html?s=76c1f7cf90754a7e33f75952ac17fcore), red letters denote a position with high matrix conservation and underlined letters indicate the mutated nucleotides; SNP affecting binding.
tein (gttTGCCAaatgtcatt) for 99307482, carbo-
hydrate response elements, made of two E-box motifs separated by 5 bp (CACGggccagcagggg), E-box binding factors (tgccCGTGgtg) for 99307484 and GATA binding factors (agcgGATAaggaa) for 99307489.

Some other important binding sites were found in both haplotypes: CCAAT/enhancer-binding protein (C/EBP) at positions 2-16 (+), 228-241 (+), 729-743 (+), 1176-1190 (+), Sp1 at 1075-1089 (-), 1133-1147 (+); cAMP response element binding protein (CREB) at 196-216 (+), 343-363 (+), 537-557 (+), 671-691 (+), 980-1000 (-); GATA binding factors at 140-152 (+), 530-542 (+), 538-550 (+); CCAAT binding factors at 194-208 (+), 842-856 (-); E-box binding factors at 305-317 (+), 634-646 (); AP2 at 957-971 (-).

The CCAAT/enhancer-binding protein (C/EBP) binding site at position 1176-1190 (+) has also been found in the bovine (Taniguchi et al., 2002), mouse (He et al., 1995) and human (Gong et al., 1996) leptin promoters. Moreover, it has been shown that in cattle (Taniguchi et al., 2002) and mouse (He et al., 1995) this element is responsible for transcriptional regulation of leptin promoter activity as well as of many adipose-specific genes (Adamowicz et al., 2006).

The Sp1 binding site at position 1075-1089 has a high degree of conservation among cattle (Taniguchi et al., 2002), mouse (He et al., 1995) and human (Gong et al., 1996); a C/G polymorphism at this site has been identified (Liefer et al., 2005) and functional analysis has been carried out (Oksanen et al., 1998) on bovine leptin expression. The Sp1 motif at 1133-1147 has also been found in man (Gong et al., 1996).

In vitro analysis of the leptin promoter

Since the in silico analysis of the promoter showed that the identified polymorphisms could affect the binding sites for many transcription factors, we decided to perform functional studies on the different promoter alleles by using in vitro gene reporter assays to test this hypothesis. To prepare phased genotypes from both parental chromosomes, we chose to amplify the promoter region of a heterozygous animal (the 99307493 SNP was not present). The PCR product was linked to the luciferase gene in the pGL3 basic vector and we obtained two constructs, named LEPF3 and LEPF46, corresponding to the in silico identified haplotype 1 and haplotype 3 leptin promoter, respectively. The LEPF3 and LEPF46 plasmids were transiently transfected into the CHO cell line to compare reporter gene activity between haplotype 1 and haplotype 3. The results of the in vitro expression analysis are reported in Figure 1.

In all the experiments, the LEPF3 construct yielded a level of luciferase expression approximately 5-8 fold above the pGL3-basic empty plasmid, while the LEP46 construct showed a decrease in transcriptional activity by approximately 2-fold compared with LEPF3. The difference observed in the gene expression results over all the experiments was statistically significant (P<0.04) and was in agreement with the in silico promoter analysis that indicated the absence of many putative transcription factor binding sites in the LEPF46 construct, with a consequent minor transcription efficiency.

In a recent work by Liu et al. (2011), a single polymorphism was identified in the pig in a more distal upstream promoter region and in an in vitro experiment using a different cell line. The authors showed a difference in luciferase activity driven by the two alleles. A previous functional study on human leptin promoter polymorphisms did not show different activities between wild-type and mutant constructs (Oksanen et al., 1998), but only a single SNP in the 5’ proximal region was considered. Moreover, in the mouse, functional in vitro studies on the leptin promoter with several mutations reported a different fold drop in activity depending on the sites of interest (Mason et al., 1997). In our case, the analyzed construct contained many mutated bases that could potentially affect the binding of more transcriptional factors, and could, therefore, lead to thinner regulation of transcriptional promoter activity. Moreover, some authors have indicated that the unique interactions of multiple SNPs within a haplotype can affect the biological phenotype while single polymorphisms within the haplotype may have poor predictive power (Drysdale et al., 2000; Winkelmann et al., 2003).

Nevertheless, it is clear that the manner and extent of leptin binding to receptors and carrier proteins determines the degree of intracellular effects; in bovine ObRub, ObRa and ObRc, spliced isoforms are all strongly expressed in intermuscular, perirenal and subcutaneous fat tissue, suggesting that these tissues respond to leptin in addition to secreting it (Kawachi et al., 2007). Various studies have shown that the ob/ob genotype in different genetic backgrounds may lead to different phenotypes (Qu et al., 2001) and that traits for leanness and the extent of fat deposition are ascribed to polygenic effects (Roehe et al., 2003).

Conclusions

In livestock production, traits such as feed intake, energy balance and fast lean growth are very important for the profitability of meat production enterprises. The identification of genetic markers that are positively associated with economically important traits has the potential to be used in breeding programs to speed up genetic improvement (Van der Lende et al., 2005). Fatness traits are important in pig production since they influence meat quality and fattening efficiency and, in addition to QTL studies, many candidate gene polymorphisms have been analyzed in terms of their associations with pig fatness (for a review see Switonski et al., 2010).

Five SNPs were previously described in small fragments of the leptin promoter region, but no evidence for an association with fatness traits was found. This work aimed to study a longer 5’ promoter region of the gene. We identified 17 new genetic variants and we report for the first time the differential transcriptional activity of two leptin pig promoter phased alleles of the same length but differing in 16 naturally occurring chromosomal positions.

Many studies have tried to find associations between the genomic context, in vitro experiments, leptin mRNA levels and in vivo measurements on animals of different breeds, but none have found clear relationships. The effects of promoter SNPs on gene transcription may be complex and involve many interactions among multiple polygenic sites. The way whereby genetic variations and polygenic effects result in variations in fat deposition via molecular regulation, in particular in different
pig breeds, is still largely undefined. The molecular study reported here may be addressed toward studies employing electrophoretic mobility shift assays (EMSA) or chromatin immunoprecipitation-sequencing (ChIP-seq) technologies to identify DNA interacting proteins directly affecting leptin gene transcription in different cellular and genetic contexts.

References

Adamowicz, T., Pisikowski, K., Starzynski, R., Zwierzchowski, L., Switonski, M., 2006. Mutation in the Sp1 motif of the bovine leptin gene affects expression. Mamm. Genome 17:77-82.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Bayerlein, M., Werner, T., 2005. Maternal factor C/EBP beta alters receptor expression and predicts body weight. Endocrinology 146:1013-1022.

Bargellini, G., Letzsch, M., Schützer, F., 2010. Leptin in farm animals: where are we and where can we go? J Appl. Genet. 51:153-168.

Benjamin, S.S., 2004. Genetic polymorphisms in porcine leptin gene and their association with reproduction and production traits. Aust. J. Agr. Res. 55:699-704.

Berezov, P., Wang, Y., Li, G., Kortt, A.A., 2006. Leptin gene polymorphisms with production traits in pigs. J. Anim. Breed. Genet. 123:378-383.

Breteler, M., Stam, H., van Duijn, C., 2000. Complex promoter and coding polymorphisms in porcine leptin gene and their association with reproduction and production traits. J. Anim. Breed. Genet. 125:703-709.

Bui, S., Pratley, R.E., Weintraub, D.B., 1996. Genomic structure and promoter analysis of the human obese gene. J. Biol. Chem. 23:3971-3974.

Chen, H., Y., Quon, M.J., Reitman, M., 1997. Characterization of the human leptin gene and their association with fatness in four pig breeds. Mamm. Genome 10:191-193.

Crisà, M., Di Clemente, G., Caserta, M., 2010. Leptin in farm animals: where are we and where can we go? Animal 5:246-267.