Genomic Structure and Promoter Analysis of the Bubalus Bubalis Leptin Gene

Francesco Napolitano*

Agricultural Research Council - Animal Production Research Centre (CRA-PCM) - Via Salaria, 31, 00015 Monterotondo (Rome), Italy

Abstract: The buffalo is an animal ever-growing in our hemisphere, one of the major problems for the genetic improvement of this species, concerns two reproductive aspects: the difficult heat detection and the seasonality of the oestrus cycles. The leptin plays a critical role in the regulation of reproductive and immune function in humans, it is at the centre of the complex networks that coordinate changes in nutritional state with many diverse aspects of mammalian biology. We have sequenced the 5' flanking region and exon 1 of the *leptin* gene in buffalo. The sequencing is the 'first step' for understanding the role of various parts of the genome and is a springboard from which to decode the genome. The simple sequencing, in fact, does not provide information directly applicable to understand the mechanisms underlying physiological and pathological processes, but represents a necessary step through which you can identify the role of different regions of DNA. However, waiting for the complete genome sequence of the buffalo, the database of the bovine genome offers the opportunity to investigate the buffalo genome in genes which are recognized to influence physiological processes related to reproduction in other species. In this context, our research had decided to investigate the *leptin* gene and particularly the regulatory area: the *promoter.*

Keywords: Single nucleotide polymorphisms, genetic markers, capillary electrophoresis.

INTRODUCTION

The buffalo is an animal ever-growing in our hemisphere, both for its greater rusticity, longevity, better adapted to humid environments, and for the highest income generated from the sale / transformation of the milk, that is not subject to quotas in European countries.

However, one of the major problems for the genetic improvement of this species, concerns two reproductive aspects: the first is the difficult heat detection, caused by silent heats, that is an obstacle to the widespread use of instrumental insemination; the second is the seasonality of the oestrus cycles, because buffaloes go into heat between October and December.

The next -generation sequencing technologies have already contributed to the characterisation of farmed animal genomes, and the development of high throughput single nucleotide polymorphisms (SNP) genotyping platforms and the first applications in animal breeding have begun to emerge. Breeding and selection of cattle, pig and sheep have started to take into account genomic information to some extent. In the next few years it is expected that also the buffalo genome will be sequenced, and it will be possible to deliver genomics-driven improvement in buffalo breeding and production.

*Address corresponding to this author at the CRA-PCM, Via Salaria 31, 00015 Monterotondo (Rome), Italy; Tel: +39 06 90090215; Fax: +39 06 9061541; E-mail: francesco.napolitano@entecra.it

In fact, the sequencing is the 'first step' for understanding the role of various parts of the genome and is a springboard from which to decode the genome. The simple sequencing, in fact, does not provide information directly applicable to understand the mechanisms underlying physiological and pathological processes, but represents a necessary step through which you can identify the role of different regions of DNA.

However, before the complete sequence of the buffalo genome is available and considering the high sequence homology between cattle and buffalo (96-97% for the coding and regulatory regions of genes), the database of the bovine genome offers the opportunity to investigate the buffalo genome in genes (candidate genes) which are recognized to influence physiological processes related to reproduction in other species. In this context, our research had decided to investigate the *leptin* gene and particularly the regulatory area: the *promoter*.

In genetics, a *promoter* is a region of DNA that facilitates the transcription of a particular gene. Promoters are located near the genes they regulate, on the same strand and typically upstream of the transcription start site. In order for the transcription to take place, the enzyme that synthesizes RNA, known as RNA-polymerase, must attach to the DNA near a gene. Promoters contain specific DNA sequences and response elements which provide a secure initial binding site for RNA polymerase and for proteins called 'transcription factors' that recruit RNA polymerase.

ISSN: 1927-5196 / E-ISSN: 1927-520X/12

These transcription factors have specific 'activator' or 'repressor' sequences of corresponding nucleotides that attach to specific promoters and regulate gene expressions.

The *leptin* gene was identified in 1995 as the product of the *obese* gene and a hormonal signal that regulates energy balance in mice. In human, Farooqi and O'Rahilly [1] have defined the role of leptin-responsive pathways in the regulation of eating behaviour, intermediary metabolism, and the onset of puberty. They also demonstrated that leptin signaling plays a critical role in the regulation of reproductive and immune function in humans, which places leptin at the centre of the complex networks that coordinate changes in nutritional state with many diverse aspects of mammalian biology.

The *leptin* gene is highly conserved across species and is located on chromosome 4q32 in the bovine [2]. Taniguchi *et al*. [3] have isolated a bovine genomic clone that contained about 3-kb in 5'-flanking region upstream from the putative transcription start site (Figure **1**).

EXPONENTIAL AMPLIFICATION BY PCR (POLYMERASE CHAIN REACTION)

Starting from the above Bos taurus sequence (Figure **1**), two amplicons were designed: red amplicon, **1141 bp** in size and blue amplicon, **613 bp** in size; so to cover the 5' flanking and exon 1 of the *leptin* gene, as indicated by Taniguchi *et al.* [4] (GenBank: AB070368).

Polymerase Chain Reaction allows us to synthesize several times (amplification) by an enzyme a specific segment of DNA located between two regions of known nucleotide sequence, producing a large number of copies through a series of reactions:

- denaturation (95°C)

- annealing of primers (40-68°C)

1 GAATTCAACA ATTCTATTTA TCAAGAAATC TCCCACAAAT ATACTCACAC TGTGCTCCAA
61 ....
.....
1381 AACATTTCCG GGGGGGGGGGG **AGGCGGAGAGGAGGAAAGAT**→TTTCTTCAA AATGTAATTT
1441 CATTGTAGAC ACTTCTTTAA AAGAAACATT TCTTTATTTG ACAGTTCCAG GCCTTAGTTT
1501 CAGCAGGCAG GATGTTTAGT CGCAGCATGA GAACTCTTAG CTGCGGCATG CGGGACCCAG
1561 TTCAGTTCCC TGACCAGATA TCGAACCTGG GGCCCCTGCA TTTGGAAGCA GGGAGTCTTA
1621 GCCACTGGAC CACCAGGGAA GTCCCCTGTA GATGTTTTTA TGAAAAGCAG AAAAGCACAA
1681 AGAAGAGCTT AAAGATTCCT GATCCTACTC CCAATAGTGA TAATGTATAT TTTGGTGTGA
1741 GAGTGTGTGT ATTGATTGGA ATGTGTGTGA TCAGAAAACA CATACCATTT TATAATCCGT
1801 TCTTTCCAGC TCACAAAATA AAGTTATTTT CCTACATCAT TAAATATTAC TTTACAACAT
1861 AATTTTTAAT GTGTGCATAT TGCTGCTATG TGATTTTCAA TAACTTACTA ATTTCCTATG
1921 CTGAACATTT AGTTGTTGTC CAACCTTTTT AGTGGCCATG TAATTATAAA TCATGGTCAA
1981 TGCTAACAAT TTCTGACCTC ACAAACATA **TAGTACAATATCCTTCCTTTCTT**→CAATAGAT
2041 AATTAT←**TAAAAGCAAAACAACCAGGC** TCAAACAAAGCAAT TATAAAATAT CTTTAAAAAG
2101 ACATTGGGTA AAATTCAAAT GCAGACTAGC TCATGATGTT AAAGAATTAC TCTTGTGTGG
2161 TAATGGTCTT GTGATAGAGA TAGAAATGCT TCCTTATTTT TCAGATAAAC ACTTAAGTAT
2221 TTAAGGATGA AACGCCCTGA TGTTTGTAAT TTGCTTTAGA ATATTTTAGC CAAAAGAATT
2281 AATGATGCAA ATATGCAAAA AGAGTACGTT AAACCTAAAT TTGCGATTTT CATTTAAAAA
2341 TATATCTTAA AAATGAAAAT CTTCGTGCAA CGCACGGGGC TATCAATGTG GGATACAGAT
2401 GTGAACAAAA CGGACCCGTG TGGGACTCGG CGGAGCACAC AGATTTTGCGGGAG **CACGTT**
2461 **CCCGTTAGGAAGTCTCTG**→ATGCAATACGAC CGGTGCCCTT CAGGACCTGT GAGACTGACT
2521 TTCC←**TTACCCCTCCACACCATCAT CAA**→GGCAGGTGTGATTTTCCAGGCCA GGCCTACGGC
2581 CGGGTTTCCC CGGGGGCCCA GAGCCGTCGG GTCTTGCCGC CCAGCGGAGC TGGCTGCTCC
2641 GGCCTCACTG TCGGGGCGCC ACCGCCCCCA GCCGGCTCAG AGGAACCCCT CACCGCCACC
2701 CTGTCCCAGG CGGCCTTTCC CCGAGGCCCG AGGGTCAGAT CCTGGGGCCA CCTCGAGGAT
2761 TTCTCACACC TGCCCAGCCA CCCCCAGCTT TTCAGGTGAT ACCGGAGGGT GGGCGTGGGG
2821 CTCCTGGCGC ATCCGAGTCC CTCCCTGGAG TCCCCGACCG CGGCCGCCCG GCCCGACGCT
2881 GCCCCGCCGC CCCGCAGGGC GGGAGCCGGC GCTGCGGGTG CGCCCCGGCC AGCCGGGCAG
2941 TTGCGCAAGT TGTGCTTCGG CGGCTATAAG AGGGGCGGGCAGGCATGGAGCCC←**CGGA(√)GGG**
3001 **ATCGAGGAATCG** CGGCGCCA GCAGCGGCGA GGTAAGTGCCCGGC←**TCTCTCCTTTCTACTT**
CTCCGCCG

Figure 1: Bos taurus gene for *leptin*, 5' flanking sequence and exon 1 GenBank: AB070368.

- polymerization of new fragments (amplicons) by Taq-Polymerase (72 °C).

The reaction is carried out in a Thermal Cycler (Figure **2**).

Figure 2: PCR – Thermal Cycler.

The amplification products are visualized by agarose gel electrophoresis (Figure **3**).

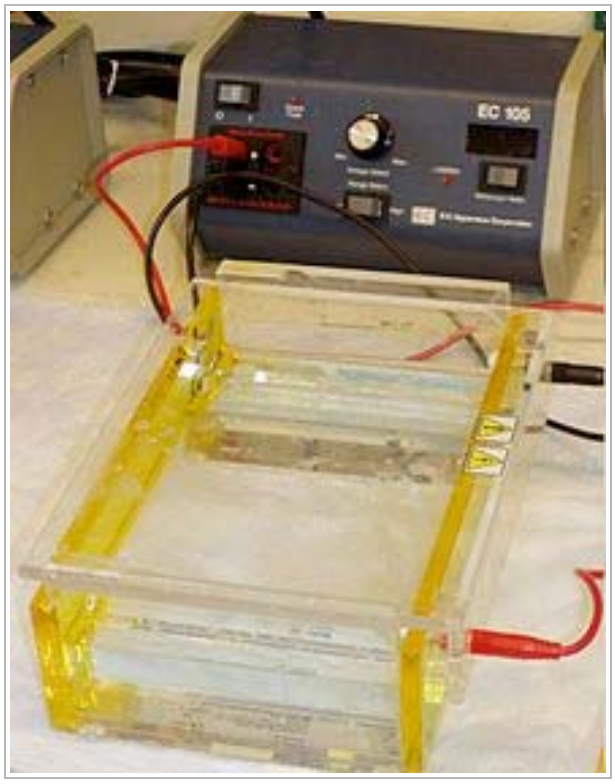

Figure 3: Gel electrophoresis apparatus – An agarose gel is placed in this buffer-filled box and electrical field is applied *via* the power supply to the rear. The negative terminal is at the far end (black wire), so DNA migrates toward the camera.

The technique allows you to view and separate nucleic acids:

DNA fragments migrate through a material selective (e.g. agarose gel) that separates them by size (molecular weight / length); smaller fragments migrate through the meshes of the gel faster than larger ones, which move more slowly.

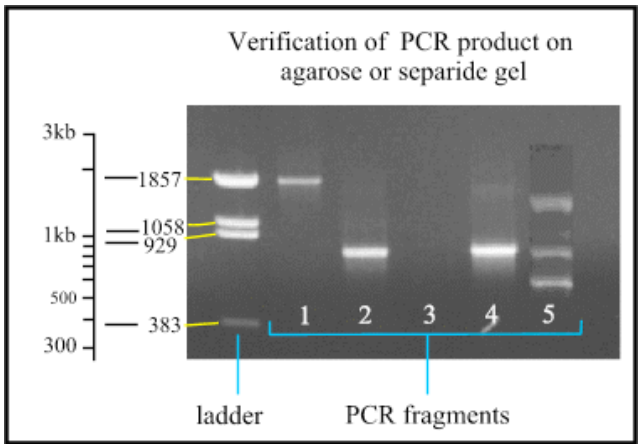

Figure 4: Running standard agarose DNA gel.

DNA of 41 non related buffaloes was PCR amplified and sequenced.

Direct sequencing by capillary electrophoresis was then performed using the Big Dye terminator v3.1 on Applied Biosystems 3500 Genetic Analyzer (Figure **5**).

CAPILLARY ELECTROPHORESIS

The principle of electrophoresis is the same: the DNA fragments migrate through a resin selective within a capillary that separates them by size (molecular weight / length), smaller fragments migrate through the meshes of the resin faster, first out from the capillary and are intercepted first by the detection system. The detection system consists of a light beam that strikes the DNA fragments.

If marked with fluorescent molecules, the DNA fragment sends an output signal that is recorded by the system.

The result is an electropherogram (Figure **6**):

The electropherogram of the same piece for more subjects allows us to obtain the complete sequence of the fragment and detected eventual variations (SNPs ↑).

SNPs are a class of molecular markers (differences due to mutations of homologous DNA regions in different individuals of the same species or different species) whose main features are:

Figure 5: Applied Biosystems 3500 Genetic Analyzer.

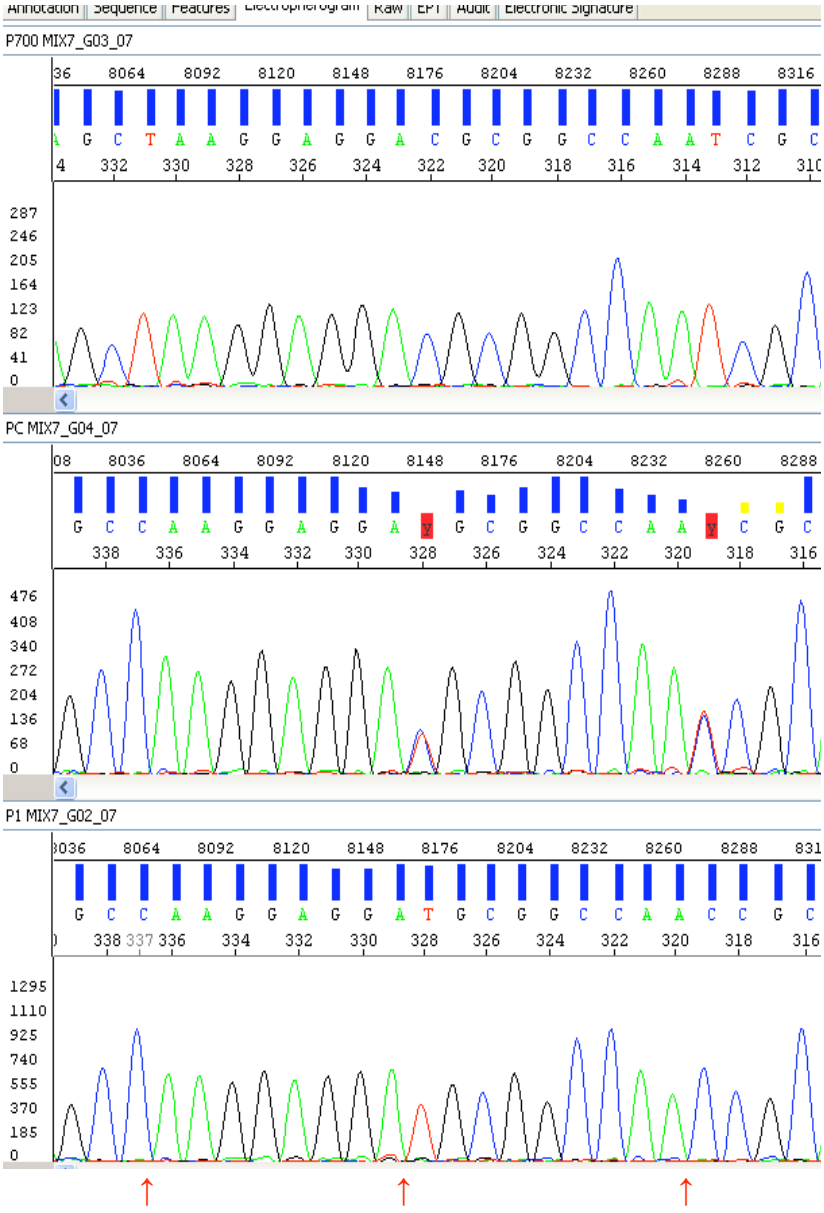

Figure 6: Electropherogram of multiple samples and SNP detection.

FEATURES Location

source	1...1691	
promoter	1...1620	
variation	83	**A** > G
variation	90	**A** > G
variation	121	**A** > G
variation	256	**G** > T
variation	283	**A** > G
variation	959	**G** > T
variation	1010	**A** > C
variation	1254	**G** > A
TATA_signal	1588 ...1592	
mRNA	1621...1654	
exon 1	1621...1654	

1 AGGCGGAGAG GAGGAAAGAT TTTCTTCAAA ATGTAATTTC ATTGTAGACA CTTCTTTAAA
61 **A**GAAACATTT CTTTATTTGA CA**A**TTCCGG**A** TCTTAGTTTC AGCACGCAGG ATGTTTAGTC
121 ACAGCATGAG AACTCTTAGC TGTGGCATGT GGGACCCAGT TCAGTTCCCT GACCAGATAT
181 CGAACCTGGG ACCCCTGCAT TTGGAAGCAG GGAGTCTTAG CCACTGGACC ACCAGGGAAG
241 TCCCCTGTAG ATGTT**G**TTAT GAAAAGCAGA AAAGCACAAA GA**A**GAGCTTA AAGATTCCTG
301 ATCCTACTCC CAATAGTGAT AATGTATATT TTGGTGTGAG AGTGTGTGTA TTGATTGGAA
361 TGTGTGGGAT CAGAAAACAC ATACCATTTT ATAATCCGGT TCTTTCCAGC TCACAAAATA
421 AAGTTATTTT CCTACATCAT TAAATATTAC TTTACAACAT AATTTTTAAT GTGTGCATAT
481 TGCTGCTATG TGATTTTCAA TAACTTACTA ATTTCCTATG CTGAACATTT AGTTGTTGTC
541 CAACCTTTTT AGTGGCCATG TAATTATAAA TCATGGTCAA TGCTAACAAT TTCTGACCTC
601 ACAAACATAT AGTACAATAT CCTTCCTTTC TTCAATAGAT AATTATTAAA AGCAAAACAA
661 CCAGGCTCAA ACAAACAAAG CAATTATAAA ATATCTTTAA AAAGACATTG GGTAAAATTC
721 AAATGCAGAC TAGCTCATGA TGTTAAAGAA TTACTCTTCT GTGGTAATGG TCTTGTGATA
781 GAGATAGAAA TGCTTCCTTA TTTTTCAGAT AAACACTGAA GTATTTAAGG ATGAAACGCC
841 ATGATATTTG TAATTTGCTT TAGAATATTT TAGCCAAAAG AATTAATGAT GCAAATATGC
901 AAAAAGAGCA CGTTAAACCT AAATTTGCGA TTTTCATTTA AAAATGTATC TTAAAAAT**G**A
961 AAATCTTCGT GCAACGCACG GGGCTATCAA TGTGGGATAC AGATGTGAA**A** AAAACGGACC
1021 CCTGTGGGAC TCGGCGGAGC ACACAGATTT TGCGGGAGCA CGTTCCCGTT AGGAAGTCTC
1081 TGATGCAATA CGACCGGTGC CTTCAGGACC TGTGAGGCTG ACTTTCCTTA CCCCTCCACA
1141 CCATCATCAA GGCAGGTGTG ATTTTCCAGG CCAGGCGTAC GGCCGGTCGC TCTCGCCAAC
1201 CCCGGGTTTC CCTGGGGGCC CAGAGCCGTC GGGTCCCGCC GCCCAGTGGA GCT**G**GCTGCT
1261 CCGGCCTCAC TGTCGGGGCG CCACCGCCCC CAGCCGGCTC AGAGGAACCC CTCACCGCCA
1321 CCCTGTCCCA GGCGGCCGTT CCCCGAGGCC CGAGGGTCAG ATCCTGGGGC CACCTCGAGG
1381 ATTTCTCATA CCTGCCCAGC CACCCCCAGC TTTTCAGGCG ATACCAGAGG GTGGGCGTGG
1441 GGCTCCTGGC GCATCCGAGT CCCTCCCTGG AGTCCCCGAC CCGCGGCCGC CCGGCCCGAC
1501 GCTGCCCCGC CGCCCCGCAG GGCGGGAGCT GGCGCTGCGG GTGCGCCCAG GCCTGCCGGG
1561 CAGTTGCGCA AGTTGTGCTT CGGCGGCTAT AAGAGGGGCA GGCAGGCATG GAACCCCGGA
1621 GGGATCGAGG AATCGCGGCG CCAGCAGCAG CGAGGTAAGT GCCCGGCTCT CTCCTTTCTA
1681 CTTCTCCGCC G

Figure 7: Bubalus bubalis leptin gene, exon 1.

- Abundance (10 million in the human genome)
- High reproducibility and accuracy
- The analysis is automated
- SNPs are often contained in the genes expressed

In this way we obtained the sequence of the *promoter* of the gene for *leptin* in Buffalo. The novel sequence (Figure **7**) was deposited with GenBank under the accession number JF681145 [5]: that showed an homology of 96% with the bovine sequence.

Eight **SNP**s were made evident within the 41 genotyped buffaloes. Under the hypothesis that genetic variants in the promoter of the leptin gene might influence the age at puberty, we have estimated the effect of each **SNP** on the variability of hormonal parameters in buffalo heifers. The preliminary results were presented in VI Buffalo Symposium of America.

ACKNOWLEDGEMENTS

This study is part of the GENZOOT research program, funded by the Italian Ministry of Agriculture (Rome, Italy).

APPENDIX

PCR: Polymerase Chain Reaction

SNP: Single Nucleotide Polymorphism

REFERENCES

[1] Farooqi IS, O'Rahilly S. Leptin: a pivotal regulator. Am J Clin Nutr 2009; 89(3 Suppl): S980-4. http://dx.doi.org/10.3945/ajcn.2008.26788C

[2] Stone RT, Kappes SM, Beattie C. The bovine homolog. Mamm Genome 1996; 7: 399-400. http://dx.doi.org/10.1007/s003359900119

[3] Taniguchi Y, Itoh T, Yamada T, Sasaki Y. Genomic structure. IUBMB Life 2002; 53: 131-5. http://dx.doi.org/10.1080/15216540211465

[4] Taniguchi Y, Itoh T, Yamada T, Sasaki Y. Bos Taurus gene for leptin, 5'flanking sequence and exon 1. GenBank: AB070368 [updated 2006 Oct 5]; Available from: http://www.ncbi.nml.nih.gov/.

[5] Scatà MC, Grandoni F, Antonelli S, De Matteis G, Moioli B, Napolitano F. Bubalus bubalis leptin gene, exon 1. GenBank: JF681145 [updated 2011 May 16]; Available from: http://www.ncbi.nml.nih.gov/.

Received on 29-03-2012 Accepted on 18-05-2012 Published on 01-09-2012

DOI: http://dx.doi.org/10.6000/1927-520X.2012.01.02.03