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.

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.
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 GAATTCACA ATTTCTATTTA TCAGAAAGAT TTCTCTTCAA AATGTAATTT
1381 AACATTTCCG GGGGGGGGGG AGGCAGAGAAGGAAAGGAT .TTTCTCTCAA AATGTAATTT
1441 CATTGAGAC ATCCTCTTTAA AAGAACACT TCTTTATTTG ATACCTTCA GCCTTAGTTT
1501 CAGCAGCGC GATGTTTAGT CGCAGCATGA GAACTCTTAG CTGCGGCATG CGGGACCCAG
1561 TTCAGTTCCC TGACACAGA TGCGACCTTG GGCCTCTGCA TTGGAAGCA GGGAGTCTTA
1621 GCCACTGCGC CACCGGGGAA GTCCCCCTTGA GATGTTTTTA TGAAAAGCAG AAAAAAGAAA
1681 AGAGAGGTCTT AAATTTTCTT ATGTTGTGTTA ATGTTGTGTTA TCAGAAACAA CATAACATT TTTACTTCTAG
1741 GATGCCTGAG TTGACCAAGA TGTTGTGTTA TCAGAAACAA CATAACATT TTTACTTCTAG
1801 TCTTTCCAGC TCACAAAAATA AAGTATAATT CCTACATCAT TAAATATATT TCCATACAC
1861 ATTTTTTTTA TGTTGCATAT TGTCTGCTGAT TGTCTTCTAA ATACCTTTA TTTCTTCTAG
1921 CTGACACATT ATGTTGTGTTA ATGACACATT TTTATATAA TTAATGCTGAA TTTACTTCTAG
1981 TGCTAAACAT TTCTACACCT CAAACACAA TAAGAAATATC GCCTTCTTTT ATTCTCTTCGT
2041 AATTAT . TAAATGAAAACAAAAACCCAGCC CAAAACAGACAT TATAATAAT CTGAAAAAAC
2101 ACATTGCGTA AATTCTATTTA GCGACTAGG TCTATGTGTT AAAAAAGA TCTCTTATCTGTGAGT
2161 TAATGTTCTT GTGATAGAGA TAGAAATGCT TCTTTATTTT CTGCAGAAG CCTAATATT
2221 TTAAGGATGAA GCACACCTCG AAAAAAGAAA TTTATACAA TTGAAACATA TTTACTTCTAG
2281 ATGAAATGCA ATATGCAAAAA AGAGACTTG TAAATATATT TTTCTTCTTT TTTACTTCTAG
2341 TATATCTTAA AATATTATG ATCTGCTGCA AAGACGGGGC TATCCTAGT GAACACGTA
2401 GCGAATACAC GAGCCGGCGG TTTGGGATCG AAGACGGCGG GAAATATGGG GACGACTT
2461 CCGTTAGGAAGAATCTCTTG .. ATGAAATGGA CAGGAGCTTCA GTGATTGACT
2521 TTTCC . TTTACCCACACACATCAT CAA .. GCCAGGTGTAAT TTCTTTTACGCA GCCGCTGCC
2581 CGGCTTCTCC CGGAGGGCCCA GACGCCTGCA GTCTTGCCGC CAGCGGAAGG TGCCGACTCC
2641 GCCCTCAGTG TCAGGGCCCG ACCGCCCCCA CGGAGCCGTA AGAAGGCCCG ACCGCCCCCA
2701 CTGCAGGAGGT GCAAACACCGG CAGAGCGCTG CAGAGCGCTG CAGAGCGCTG
2761 TTCTGAGAGG GAAACCCACG CCCCCCGCTT ATCCAGGCGA ATCGGAACCG CAACCTCGG
2821 TAAGAATACAC GAGCCGGCGG TTTGGGATCG AAGACGGCGG GAAATATGGG GACGACTT
2881 CGGCTTCTCC CGGAGGGCCCA GACGCCTGCA GTCTTGCCGC CAGCGGAAGG TGCCGACTCC
2941 GTGATAGAGA TAGAAATGCT TCTTTATTTT CTGCAGAAG CCTAATATT
3001 ATTCAGAATCG CCGAGCCCA GCAGGGCGGA GTGAAATACAC CACCGGCCCG CGGGCCCGCC

*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.](image)

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.](image)

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.](image)

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):

![Figure 6: Electropherogram](image)

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.
<table>
<thead>
<tr>
<th>FEATURES</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>1...1691</td>
</tr>
<tr>
<td>promoter</td>
<td>1...1620</td>
</tr>
<tr>
<td>variation</td>
<td>83 (A &gt; G)</td>
</tr>
<tr>
<td>variation</td>
<td>90 (A &gt; G)</td>
</tr>
<tr>
<td>variation</td>
<td>121 (A &gt; G)</td>
</tr>
<tr>
<td>variation</td>
<td>256 (G &gt; T)</td>
</tr>
<tr>
<td>variation</td>
<td>283 (A &gt; G)</td>
</tr>
<tr>
<td>variation</td>
<td>959 (A &gt; C)</td>
</tr>
<tr>
<td>variation</td>
<td>1210 (A &gt; G)</td>
</tr>
<tr>
<td>variation</td>
<td>1254 (G &gt; A)</td>
</tr>
<tr>
<td>TATA_signal</td>
<td>1588...1592</td>
</tr>
<tr>
<td>mRNA</td>
<td>1621...1654</td>
</tr>
<tr>
<td>exon 1</td>
<td>1621...1654</td>
</tr>
</tbody>
</table>

```
1 AGGCCGAGAG GAGGAAGAT TTCTTCAA ATGTAATTC ATGGTAGACA CTCTTTAAA
61 AGAAACATT CTTATTTGGA CAAATTCGGAG TCCTTTTC GAGATTTAAT TCTTCAATTC
121 AAGCCATGAG AACTCTTCGG TGGGCAATGT GGGACCCAGT TCAATTCCC TTCCCAATAT
181 CGAACCTGGG ACCCCCTGCTT TGGGAACGAG GAGCTTCTT TCCACTACGT ACAAGAGTAT
241 TCCCCCTGAG ATGTTTCTT TAAAGACAGA AAAGCAGAA GAGAGCTTTA AAGATCT CCTT
301 ATCTCTACTC CAATGTTTAC ATGTTTTGAG AGTGTGTTG GATTTTGGGCAA
361 TGTTGTTTGG GAGAAGACAT ATACCACTTT TAAATCCTGG TCTTCCCAGC TGCAAAAAAT
421 AGGTTATTTC CACATCATT TAAATTTGAG ATTTCAAGAT TTCAACTAAT GTTTTGGCTTAC
481 TCCTGCTATG TGAATTTTTCA TAAATTTTCTT TCTTCTTCTT GTGGTAAAGTG TTCTTGA
541 CAAACCTTTTT ATGTTTGGCA TAAATAAAA TCTTGTGCAA TGTTCAAAAT TCTTTGACCTCTC
601 ACAAAATAAT AGTACAAAT ATCTTCTTT TCAAAAGAT ATTTAAAGAA ACAGAAAAAC
661 CCACGGCTCAA ACAAAACAGA CAATTAAAA ATATCTTTTT AAAAGACATG GGTAAAATTTCC
721 AAATGGACAG CTACTCTAGA GGTAAAAAG ATTTTAAATATTAATAGA ATTTTTTTAT
781 GAGATTAAAA GAGTTTCTTT TTTTTCAAT AACACTGAAG GAGTTTTAGG ATGAAAAAC
841 ATGATTATTTG TTAAATTTCT TAGAATATT TAGCCTAAAAT AATTATGAT CAAATAATGC
901 AAAAAACGCA CGTTAAACCT AATTGTCGTA TTTTTTACCA AAAATGTATG TTTAAAAATGGA
961 AAATCTTCTG CCAACTGAGAC GGGACTACAA TTCTGTTGAC ATGTTTGGAA AAAACGGAC
1021 CCTGTGAGGAC TCGGCGGAGC ACACAGAATT TGGGAGGAGA CTGTCCCTGGT TGGAGTCTC
1081 TGATGCAATA CGAAGGAGCTG TCTTGGGACT TCTTCTTTT TCTTCTTTCT
1141 CCACAGCTCA CCAAGATCTG TTCTTTGAG TACAGAGTGA AAAAAAAAC
1201 CCGGCTTCA CCGGCTTCGG GAGGAAAGAC TGTGTTGTAG AAAAAAAGAC
1261 CCAGCTTTCG CCAAGGCTCA CCGGCTTTC GAGGAAAGAC TGTGTTGTAG AAAAAAAGAC
1321 TCTTCGCTGG CACAGGGCTG TCGGCTGCTG ATGTTTGGGA AAAAAAAGAC
1381 ATTCACAAAT TGGGACAGG ATGTTTGGGA AAAAAAAGAC
1441 GCCGCTTGAG GAGGAAAGAC TGTGTTGTAG AAAAAAAGAC
1501 GTCCCTTCCG CAGCCCTTCCGC CAACTGAGAC TGTGTTGTAG AAAAAAAGAC
1561 CAGATTGCAA ATGTTGCTG TCACAGGCTG TCGGCTGCTC ACGCCCTTCC
1621 GGATCTGGAG ATTCGCGGGC CAGGAGCTGG TACAGGGCTG TACAGGGCTG
1681 CTCTCTGGCC ACAGCCTTGG CAGGAGCTGG TACAGGGCTG TACAGGGCTG
```

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 SNPs 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

   http://dx.doi.org/10.3945/ajcn.2008.26788C

   http://dx.doi.org/10.1007/s003359900119

   http://dx.doi.org/10.1080/15216540211465


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

© 2012 Francesco Napolitano; Licensee Lifescience Global.
This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License
(http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in
any medium, provided the work is properly cited.