

Genomic Organization and Genetic Mapping of the Bovine PREF-1 Gene

Scott C. Fahrenkrug,¹ Brad A. Freking, and Timothy P. L. Smith

U.S. Department of Agriculture, Agriculture Research Service, U.S. Meat Animal Research Center, Clay Center, Nebraska 68933-0166

Received September 13, 1999

As a potential regulator of nutrient partitioning in beef cattle, we have cloned and genetically mapped the bovine PREF-1 gene. A full-length PREF-1 cDNA was isolated by iterative purification from a mixed-tissue cDNA library to which adipose contributed mRNA. Analysis of partial cDNAs from this library revealed that the 3'-terminal exon of the bovine PREF-1 mRNA is spliced in a manner analogous to its murine ortholog. However, we have also detected a PREF-1 splice form apparently unique to cattle. Aside from this alternative selection of a splice donor in the bovine fifth exon, the exon/intron junctions of the bovine PREF-1 gene recapitulate those observed for mice. The sequences proximal to the bovine PREF-1 transcription start site are homologous to the mouse PREF-1 promoter. Importantly, the sequence experimentally identified as critical to PREF-1 "suppression in adipocyte differentiation" is conserved in the bovine gene. The bovine PREF-1 gene was mapped to the telomeric end of BTA 21 by virtue of a physically linked microsatellite with seven alleles and 285 informative meioses. © 1999 Academic Press

The adipocyte serves as a storage depot for excess energy reserves in the form of fatty acids and triacylglycerides. These reserves are mobilized when caloric expenditure surpasses intake. Though fat serves a crucial function in the storage, metabolism and release of lipids, the diversion of feed into adipose production results in significant loss to meat producers. In addition, health-conscious consumers desire a low-fat, yet tender and flavorful beef product. Genetic variation for carcass and meat palatability traits has been described among and within cattle populations [1, 2]. Dissection of the underlying genetic variation responsible would

permit an evaluation of carcass merit early in the life-cycle.

The ontogeny of adipocytes is not well understood, though they are known to arise from mesenchymal precursors widely distributed in the vertebrate embryo [3]. One of the molecules involved in controlling adipocyte differentiation is the epidermal growth factor (EGF)-like homeotic protein *pref-1* (a.k.a. *dlk*, *zog-1*, *FA-1*, *SCP-1* [4–9]). The expression of PREF-1 is suppressed during *in vitro* preadipocyte differentiation. Adipogenic conversion requires PREF-1 suppression since ectopic expression inhibits *in vitro* preadipocyte differentiation [10, 11]. PREF-1 is also expressed in fetal pancreatic and adrenal tissues, becoming restricted to cells of the adrenal medulla, the zona glomerulosa, pancreatic β -cells and the somatotroph cells of the pituitary gland later in development [6, 12, 13]. Thus, PREF-1 expression may play an important role in the development, differentiation and function of tissues central to the endocrine regulation of nutrition partitioning.

The protein encoded by the PREF-1 gene is a transmembrane molecule comprised of six extracellular, cysteine-rich, EGF-like repeats and a short cytoplasmic tail, similar to the *Drosophila* neurogenic protein Delta [7, 14]. Alternative-splicing of the mouse PREF-1 transcript has been observed that gives rise to four splice-forms, generating proteins with various in-frame deletions in the terminal EGF-like repeat and juxtamembrane domain [14]. The *pref-1* proteins encoded by the two largest splice-forms undergo cleavage in the juxtamembrane region to release a soluble 50-kDa ectodomain [15], raising the possibility that *pref-1* may influence not only the development and/or function of endocrine tissues, but may itself act as a hormone.

We report the genomic structure of the bovine PREF-1 gene which will facilitate the assessment of allelic variation for this gene in beef cattle populations. We also report the development of a polymorphic marker closely linked to this gene, which will allow for the assessment of this gene as a positional candidate

¹ To whom correspondence should be addressed at USDA, ARS, U.S. Meat Animal Research Center, Spur 18D, P.O. Box 166, Clay Center, NE 68933-0166. Fax: 402-762-4390. E-mail: fahrenkrug@email.marc.usda.gov.

```

1   GGAGAGCGCAGCGCGCAGCCCCGGTGCAGCCCTGGCTTTCCTCGCCGCGCGCCCGCGCC   60
61  CCCTTTCGCGTCCGCAACCAGAAGCCCAGCGCGCGCCCGGAGCAGGTCCCAGCGCCCGCG   120
121 CCGTCCCCTGGGACGCGCGCCCCCGCGCCCGCGATGGCCGCGACCGCAGCCCTCTCTGC   180
      M A A T A A L L
181 CCGCCCTCTTGTCTCTGCTGGCTTTCGCGCCGAGTGCATGGAGCTGAATGCTTCCCAGG   240
      P A L L L L L A F G R S A H G A E C F P
241 CCTGCCACCCTGAAAATGGATTCTGCGACGATGACAGTGTGTGCAGGTGCCAGCCTGGCT   300
      A C H P E N G F C D D D S V C R C Q P G
301 GGCAGGGTCCCCTGTGTGACCACTGCGTGCACCTTTCCTCGGCTGTGTGAACGGCCTCTCGG   360
      W Q G P L C D Q C V T F P G C V N G L C
361 TGGAGCCATGGCAGTGCATCTGCAAGGACGGCTGGGACGGACACCTCTGTGACCTAGACA   420
      V E P W Q C I C K D G W D G H L C D L D
421 TCCGGGCTTGACCTCGACCCCTGCGCCAACAACGGCACCTGCCTGAACCTCGATGACG   480
      I R A C T S T P C A N N G T C L N L D D
481 GCCAGTACGAGTGTCTCTGCGCCCCCGGTTCTCAGGAAAGGATTGTGAGAAATGGATG   540
      G Q Y E C S C A P G F S G K D C Q E M D
541 GGCCCTGCGTGGTGAATGGCTCGCCCTGCCAGCACGGAGGAGCTGCGTGGACGATGAGG   600
      G P C V V N G S P C Q H G G S C V D D E
601 GCCGGGCCCCCACGCTGTCTGCTGTGCCCCCTGGCTTCTCGGGCAACTTCTGCGAGA   660
      G R A P H A V C L C P P G F S G N F C E
661 TCGTGACCAACAGCTGCATCCCCAACCCGTGCGAGAACCAGGGCATCTGCACCGACATCG   720
      I V T N S C I P N P C E N Q G I C T D I
721 GGGGTGACTTCCGCTGCGGTTGCCCGCGGCTTCATGGACAAGACCTGCAGCCGCCCGG   780
      G G D F R C R C P A G F M D K T C S R P
781 TGAACACCTGCACCAGCGAGCCGTGCCTCAACGGCGGCACCTGCCTGCAGCACTCCCAGG   840
      V N T C T S E P C L N G G T C L Q H S Q
841 CCATCTGCTTACCATCTCTGGGCGTGTCAACAGCCTGGTGGTCTCTGGGCACCATGGGCA   900
      A I C F T I L G V L T S L V V L G T M G
901 TCGTCTTCTCAACAAGTGCAGGCGCTGGGTGTCCAATCTGCGCTACAACCACATGTTGC   960
      I V F L N K C E A W V S N L R Y N H M L
961 GCAAGAAGAAGAACTGCTGCTGCACTACAACAGCGGGGAGGAGCTGGCCGTCAACATCG   1020
      R K K K N L L L H Y N S G E E L A V N I
1021 TCTTCCCGGAGAAGATCGACATGACCACCTTACCAAGGAGGCGCGGAGGAGAGATCT   1080
      V F P E K I D M T T F T K E A G E E E I
1081 GAGCAGCGCCCCACCGCCCCCCTTCTCGGGGTCCCAGAGCCCCCTGTCTCTCTG   1140
1141 TGCGGTCTGTTCTTATCTTTGTGGTGAATTTGCTCTCCTTTGTGTCAAATCTGGTGAAC   1200
1201 GCTACGCTCCCCTCTCTGCCTTTGCGCTGCCGTGTGTGCAACCAACGTAATTGCCAG   1260
1261 ATGAATCCTCTTCTCTTCTTAATGCATGATATAAAAAATAATAATAACGAT

```

FIG. 1. The sequence of the isolated full-length bovine PREF-1 cDNA is presented. The initiation and termination codons are indicated in bold. The amino acid sequence of the pref-1 protein is presented below the coding sequence. The putative signal peptide is italicized, the predicted EGF repeats are underlined, the predicted transmembrane domain is italicized and underlined, the potential N-glycosylation sites are underlined and bold.

for production trait quantitative trait loci (QTL) in cattle.

MATERIALS AND METHODS

PCR primers. PCR primer pairs were designed using Oligo 5.0 (NBI) and synthesized on an Oligo 1000M oligonucleotide synthesizer (Beckman). Amplification primers were designed from partial cDNA sequence of the cattle PREF-1 gene (AB009278).

Isolation of cDNA clones. The cDNA clones were isolated by an iterative screen [16] with PREF-1 specific primers (CCATGGG-CATCTCTCCTCA-3' and 5'-GCCGGCCTCCTTGGTGAA-3') from a bovine library prepared using pooled RNA isolated from mesenteric lymph node, hilar lymph node, ovary, kidney-pelvic-heart fat, hypothalamus, pituitary and subcutaneous fat (Smith *et al.*, in preparation). Alternative splice forms were detected by amplification with the primers 5'-GGTGACTTCCGCTGCCGTGTC-3' and 5'-CTACGGGTACC-ACGGGTCCTG-3' which bracket the alternative splice-donors and splice-acceptor. PCR products were cloned using the TOPO-TA cloning

```

A. /AGACCTGCAGCCGCCCGGTGAACACCTGCACCAGCGAGCCGTGCCTCAACGGCGGCACCTGCCT
C2. /AGACCTGCAGCCGCCCGGTGAACACCTGCACCAGCGAGCCGTGCCTCAACGGCGGCACCTGCCT
E. /AGACCTGCAGCCGCCCGGTGAACACCTGCACCAGCGAGCCGTGCCTCAACGGCGGCACCTGCCT
   K T C S R P V N T C T S E P C L N G G T C L

A. GCAGCACTCCCAGGTGAGCTTCGAGTGTCTGTGCAAGCCCGGTTACCCGGCCCCCGGTGTGGC
C2. GCAGCACTCCCAG-----
E. GCAGCACTCCCAGGTGAGCTTCGAGTGTCTGTGCAAGCCCGGTTACCCGGCCCCCGGTGTGGC
   Q H S Q V S F E C L C K P A F T G P R C G

A. CGGAAGCGCGCGGGCGGGCCCCAGCAGGTACCCCGTCTGCCAGCGGTTACGGGCTGACCTACC
C2. CGGAAGCGCGCGGGCGGGCCCCAGCAG-----
E. -----
   R K R A A G P Q Q V T R L P S G Y G L T Y

A. GCCTGACCCCGGGGTGCACGAGCTGCCGGTGCCGAGCCCGAGCACCCGCTCTGAAGGTGTCT
C2. -----
E. -----
   R L T P G V H E L P V P Q P E H R V L K V S

A. CATGAAGGAGCTCAACAAGAGCACTCCGCTCCTCTCCGAGGGACAGGCCATCTGCTTCACCATC/
C2. -----GCCATCTGCTTCACCATC/
E. -----GCCATCTGCTTCACCATC/
   M K E L N K S T P L L S E G Q A I C F T I

```

FIG. 2. Alignment of partial cDNAs representing alternative splice-forms of the bovine PREF-1 mRNA. The dashed line indicates that the clone does not contain the sequence presented. The amino acid sequence of the peptide encoded by the partial cDNAs is indicated in bold. The sixth EGF repeat is underlined. The potential N-glycosylation site present exclusively in splice form A is italicized.

kit (Invitrogen), identified by colony hybridization and sequenced. Comparison of bovine PREF-1 to mouse and human PREF-1 sequences were conducted using the BLAST algorithm [17, 18] with equivalent portions of cDNAs in the case of alternative splice-forms.

Isolation of genomic clones. The PREF-1 cDNA was used as a probe to screen a bovine BAC library (RPCI-42) by high-density filter hybridization (Warren *et al.*, in preparation). A total of 10 BACs were identified and isolated. DNA was pooled, digested by *Pst*I or *Sau*3A I (LTI), and subcloned. Colonies containing portions of the PREF-1 gene were identified by colony lift hybridization. Clones were isolated and sequenced with cDNA-specific and vector primers on an ABI 377 automated DNA sequencer (PE Applied Biosystems, Foster City, CA). Sequences were aligned using VectorNTI (Informax).

Marker development. Large insert genomic clones were subcloned by standard techniques [19]. Microsatellites were identified and sequenced as described [20]. A PREF-1 linked microsatellite was identified and the primers 5'-GCAGAGCAGGTGCAGGAC-3' and 5'-CCACCACTCTCCTCTCCCTTA-3' were developed for genotyping.

Linkage analysis. Marker genotypes were generated from bovine reference populations and entered into the U.S. Meat Animal Research Center (MARC) relational database [21-23]. Linkage analysis were performed using Cri-Map version 2.4 software [24].

RESULTS AND DISCUSSION

Cloning of the PREF-1 cDNA. We isolated bovine a PREF-1 cDNA via iterative screen of the MARC bovine cDNA library #1 (Smith *et al.*, in preparation) using primers based on a partial bovine PREF-1 cDNA (GenBank AB009278 by Minoshima *et al.*, 1997). A 176-bp product was generated by PCR amplification of cDNA and its identity was verified by sequencing. The clone with the longest 5'-end was identified by amplification with the SP6 vector-specific primer and a reverse PREF-1 specific primer. After four rounds of iterative

screening a single 1315-bp cDNA clone was isolated (Fig. 1, AF181462). The isolated cDNA extends 34-bp further 5-prime than AB009278 and used the same polyadenylation site, though no consensus poly(A) signal is present. Comparison of the bovine PREF-1 cDNA with that of other species reveals 87% identity with human and mouse sequences. An analysis of the predicted amino acid sequence of bovine pref-1 reveals approximately 82 and 80% amino acid identity with human and mouse pref-1 proteins, respectively. Though most of the protein domains previously identified [14] as important for pref-1 function are encoded in the isolated bovine cDNA (Fig. 1), like AB009278, our bovine PREF-1 cDNA does not encode the entire sixth EGF repeat, the entire juxtamembrane domain, or the third potential N-glycosylation site. This is analogous to splice-form C2, previously described for mouse PREF-1 [14]. This splice-form of mouse PREF-1 is predicted to encode a form of the protein that is not susceptible to cleavage, and so does not give rise to soluble pref-1 [15]. This would have important implications for the potential function of the pref-1 protein as a soluble signaling molecule. It has been suggested that membrane and secreted pref-1 protein variants play opposite roles in the control of adipogenesis and that the pref-1 protein not only participates in processes leading to inhibition of adipogenesis, but that the control of its expression and different spliced variants is essential for the adipogenic response to extracellular signals [25]. We undertook the characterization of additional cDNAs present in all four MARC bovine cDNA librar-

Genomic organization of the PREF-1 gene. Sequence data derived from genomic subclones indicated that the bovine PREF-1 gene is comprised of six exons (AF181463, AF181464, AF181465, AF181466) like its murine counterpart. The intron/exon boundaries are indicated in Fig. 3. The phase with respect to the open reading frame and the position of the splice-junctions is completely conserved between mice and cattle, indicating the isolation of the bovine PREF-1 orthologue [14]. As mentioned previously, the 3'-end of the bovine PREF-1 cDNA does not reveal any obvious polyadenylation site. However, because the genomic sequence at the 3'-end of this gene does not contain an A-rich sequence, the termination site detected is likely to be legitimate, suggesting a non-canonical poly(A) signal may be used. Comparison of bovine PREF-1 genomic sequence with the mouse PREF-1 cDNA (D16847) reveals sequence extending 84 bp beyond the detected bovine transcriptional termination site that is 89% identical to the 3'-end of the mouse PREF-1 cDNA, suggesting the possibility that the bovine PREF-1 mRNA may in some cases be polyadenylated at an analogous position. A more extensive analysis is needed to determine whether alternative polyadenylation may play a role in the regulation of PREF-1 expression.

We have also isolated a genomic clone containing sequence proximal to the detected bovine PREF-1 start-site (Fig. 3, AF181831). Comparison of this sequence to the published portion of the mouse PREF-1 promoter suggests our cDNA is nearly full-length, with a CAP site homologous to that for mouse only 17 nt upstream [14] (Fig. 3). The full-extent to which the bovine and murine promoters are conserved cannot be assessed because the full murine PREF-1 promoter sequence is unpublished. However, a bovine sequence element resembling that identified as a "suppression in adipocyte differentiation" (SAD) element in the mouse PREF-1 gene [26], is located at approximately the same distance from the transcription start site (-154 to -193 for mouse, -156 to -193 for bovine, Fig. 3). This suggests the importance of this element and a conservation of the mechanism employed for regulating PREF-1 gene expression.

Mapping of the bovine PREF-1 gene. A number of QTL affecting fat deposition in cattle have been detected [27]. To determine whether bovine PREF-1 gene is in a chromosomal position potentially important to meat quality traits we developed a polymorphic marker to position it on the bovine linkage map. The human PREF-1 gene has been physically mapped, using the GB4 cR3000 RH panel, to the telomeric end of HSA 14 [28]. The mouse PREF-1 gene has been reported as residing on the mouse X-chromosome [29]. However, this assignment is unlikely based on the observation that a genetic knockout of the mouse

PREF-1 gene does not segregate in sex chromosome-linked manner (Smas, personal communication, 1999). The map position of bovine PREF-1 was determined by linkage to a microsatellite (AF181714) isolated from a bovine BAC containing PREF-1. Genotyping of the microsatellite locus in MARC bovine reference families identified seven alleles and produced 285 informative meioses. Linkage analysis positioned the marker at the telomeric region of BTA 21. Two-point linkage analysis revealed PREF-1 was linked (LOD > 3.0) to 23 markers previously mapped to BTA 21 with the greatest support (LOD = 54.91) for linkage to BMS743 at a recombination frequency of 0.03. Multipoint linkage analysis suggested the most likely interval to be between IDVGA-30 and OY3 (<http://sol.marc.usda.gov>). These data support the predicted conservation of synteny between the telomeric regions of HSA 14 and BTA 21 [30].

ACKNOWLEDGMENTS

We thank Jeni Kuszak for technical assistance and Sherry Kliver for manuscript preparation.

REFERENCES

1. Koch, R. M., Cundiff, L. V., and Gregory, K. E. (1982) *J. Anim. Sci.* **55**, 1319-1329.
2. Wheeler, T. L., Cundiff, L. V., Koch, R. M., and Crouse, J. D. (1996) *J. Anim. Sci.* **74**, 1023-1035.
3. Lane, M. D., Tang, Q.-Q., and Jiang, M.-S. (1999) *J. Anim. Sci.* **77**, 23-32.
4. Bachmann, E., Krogh, T. N., Hojrup, P., Skjodt, K., and Teisner, B. (1996) *J. Reprod. Fertil.* **107**, 279-285.
5. Halder, S. K., Takemori, H., Hatano, O., Nonaka, Y., Wada, A., and Okamoto, M. (1998) *Endocrinology* **139**, 3316-3328.
6. Jensen, C. H., Krogh, T. N., Hojrup, P., Clausen, P. P., Skjodt, K., Larsson, L. I., Enghild, J. J., and Teisner, B. (1994) *Eur. J. Biochem.* **225**, 83-92.
7. Laborda, J., Sausville, E. A., Hoffman, T., and Notario, V. (1993) *J. Biol. Chem.* **268**, 3817-3820.
8. Okamoto, M., Takemori, H., Halder, S. K., Nonaka, Y., and Hatano, O. (1998) *Endocr. Res.* **24**, 515-520.
9. Smas, C. M., and Sul, H. S. (1993) *Cell* **73**, 725-734.
10. Smas, C. M., and Sul, H. S. (1996) *Int. J. Obes. Relat. Metab. Disord.* **20**, S65-S72.
11. Smas, C. M., and Sul, H. S. (1997) *Crit. Rev. Eukaryot. Gene Expr.* **7**, 281-298.
12. Jensen, C. H., Teisner, B., Hojrup, P., Rasmussen, H. B., Madsen, O. D., Nielsen, B., and Skjodt, K. (1993) *Hum. Reprod.* **8**, 635-641.
13. Tornehave, D., Jensen, C. H., Teisner, B., and Larsson, L. I. (1996) *Histochem. Cell Biol.* **106**, 535-542.
14. Smas, C. M., Green, D., and Sul, H. S. (1994) *Biochemistry* **33**, 9257-9265.
15. Smas, C. M., Chen, L., and Sul, H. S. (1997) *Mol. Cell. Biol.* **17**, 977-988.
16. Sonstegard, T. S., Lopez-Corrales, N. L., Kappes, S. M., Beattie, C. W., and Smith, T. P. (1997) *Mamm. Genome* **8**, 751-755.

17. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
18. Tatusova, T. A., and Madden, T. L. (1999) *FEMS Microbiol. Lett.* **174**, 247–250.
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Cold Spring Harbor Laboratory Press., Cold Spring Harbor, New York.
20. Stone, R. T., Pulido, J. C., Duyk, G. M., Kappes, S. M., Keele, J. W., and Beattie, C. W. (1995) *Mamm. Genome* **6**, 714–724.
21. Kappes, S. M., Keele, J. W., Stone, R. T., McGraw, R. A., Sonstegard, T. S., Smith, T. P., Lopez-Corrales, N. L., and Beattie, C. W. (1997) *Genome Res.* **7**, 235–249.
22. Bishop, M. D., Kappes, S. M., Keele, J. W., Stone, R. T., Sunden, S. L., Hawkins, G. A., Solinas-Toldo, S., Fries, R., Grosz, M. D., Yoo, J., *et al.* (1994) *Genetics* **136**, 619–639.
23. Keele, J. W., Wray, J. E., Behrens, D. W., Rohrer, G. A., Sunden, S. L., Kappes, S. M., Bishop, M. D., Stone, R. T., Alexander, L. J., and Beattie, C. W. (1994) *J. Comput. Biol.* **1**, 65–76.
24. Green, P., Falls, K., and Crooks, S. (1990) Washington Univ. School of Medicine, St. Louis, MO.
25. Garces, C., Ruiz-Hidalgo, M. J., Bonvini, E., Goldstein, J., and Laborda, J. (1999) *Differentiation* **64**, 103–114.
26. Smas, C. M., Kachinskas, D., Liu, C. M., Xie, X., Dircks, L. K., and Sul, H. S. (1998) *J. Biol. Chem.* **273**, 31751–31758.
27. Casas, E., Shackelford, S. D., Keele, J. W., Stone, R. T., Kappes, S. M., and Koohmaraie, M. (2000) *J. Anim. Sci.* **78**, in press.
28. Deloukas, P., Schuler, G. D., Gyapay, G., Beasley, E. M., Soderlund, C., Rodriguez-Tome, P., Hui, L., Matise, T. C., McKusick, K. B., Beckmann, J. S., *et al.* (1998) *Science* **282**, 744–746.
29. Brady, K. P., Rowe, L. B., Her, H., Stevens, T. J., Eppig, J., Sussman, D. J., Sikela, J., and Beier, D. R. (1997) *Genome Res.* **7**, 1085–1093.
30. Chowdhary, B. P., Fronicke, L., Gustavsson, I., and Scherthan, H. (1996) *Mamm. Genome* **7**, 297–302.