ABSTRACT

Osteopontin (OPN) is a highly phosphorylated glycoprotein whose gene has been cloned and sequenced in different species. Several whole genome scans have identified quantitative trait loci (QTL) affecting milk production traits on bovine chromosome 6 close to the osteopontin gene (OPN) location. The presence of OPN in milk and its elevated expression in mammary gland epithelial cells together with previous QTL studies have prompted us to investigate the effects of OPN variants on milk production traits in the Holstein dairy cattle population. A single nucleotide polymorphism in intron 4 (C/T) was detected and primers were designed to amplify genomic DNA from 1362 bulls obtained from Cooperative Dairy DNA Repository and from 214 cows from the University of Wisconsin herd. For the Repository population, the C allele was associated with an increase in milk protein percentage and milk fat percentage. Correlation between milk protein percentage and milk fat percentage was about 0.57. For the University of Wisconsin herd, the estimates of the effects of allele C were in the same direction as for the Repository population, although these estimates did not reach statistical significance. Our results are consistent with other studies that showed a significant association of the microsatellite markers in the region of OPN with milk protein percentage and other correlated traits. (Key words: osteopontin, quantitative trait loci, production trait)

INTRODUCTION

Osteopontin (OPN) is a highly phosphorylated glycoprotein whose gene (OPN) has been cloned and sequenced in different species. Comparative sequence analysis of the bovine OPN cDNA in various species has revealed both conserved and nonconserved sequences (Kerr et al., 1991). It was found, for example, that the bovine and ovine sequences have a 22-AA gap compared with all other examined species. Bovine OPN consists of 6 exons spanning about 7 kb of genomic DNA (GenBank accession number: NW_255516) and encodes a 278-AA protein (Kerr et al., 1991). Since its first description in 1979 as a protein associated with malignant transformation, OPN has been intensively studied in human, mouse, and sheep. It has been suggested that human OPN has various roles in cell adhesion, chemotaxis, cell survival, tissue remodeling, regulation of inflammation, fetal growth and development, and in initiating and maintaining pregnancy (Denhardt et al., 2001; Johnson et al., 2003).

Constitutive expression of OPN exists in several tissues, and the protein is present in milk, plasma, and urine. The OPN concentration in human milk ranges from 3 to 10 μg/mL (Senger et al., 1989). Using microarray analysis of RNA from human milk cells, Nagatomo et al. (2004) found that OPN showed the highest expression among 240 genes examined. They also found that both mRNA and protein levels were highly expressed throughout the entire lactation. The presence of OPN in milk and the high expression in mammary gland epithelial cells may account for the proliferation and differentiation of mammary glands (Nagatomo et al., 2004). The major sources of OPN were mammary gland epithelial cells and monocytes and macrophages in milk. Osteopontin has also been detected in raw milk of cows at a concentration of 8 mg/L (Bayless et al., 1997). This has prompted us to investigate the effects of OPN on milk production traits in dairy cattle.

Previously, several whole genome scans have identified QTL affecting milk production traits on bovine chromosome 6 close to the OPN location (Zhang et al., 1998; Mosig et al., 2001; Nadesalingam et al., 2001; Ron et al., 2001; Rodriguez-Zas et al., 2002; Ashwell et al., 2004; Olsen et al., 2004). Ron and colleagues (2001) localized a QTL affecting protein percentage to a confidence interval of 4 cM in the region of OPN. Based on the aforementioned studies on the expression of OPN in the mammary gland and milk production QTL near...
the gene, we investigated possible associations between variants of the gene and milk production traits in Holstein dairy cattle.

**MATERIALS AND METHODS**

**Data**

Semen samples from 28 Holstein sires and their 1362 sons (19 to 102 sons per sire) were obtained from the Cooperative Dairy DNA Repository (CDDR), which is maintained by the USDA Bovine Functional Genomics Laboratory (Beltsville, MD). In addition, 214 blood samples were obtained from the herd of the University of Wisconsin (UW). Predicted transmitting abilities data for milk yield, milk protein and fat yields, milk protein and fat percentages, and SCS were obtained from the Animal Improvement Programs Laboratory (Beltsville, MD). Summary statistics of PTA of both sons from the CDDR sire families and of cows from the UW herd for production and health traits is given in Table 1.

**Genotyping**

Genomic DNA was extracted from semen samples using proteinase K and phenol/chloroform according to the procedures of Kappes et al. (2000) and from blood samples using GFX Genomic Blood DNA Purification kit (Amersham Biosciences, Piscataway, NJ). The DNA concentration was measured using a spectrophotometer (Ultraspex 2100; Amersham Biosciences). A total of 1604 samples were genotyped in this study: 28 sires and their 1362 sons and 214 cows of the UW herd. To detect single nucleotide polymorphisms (SNP) in OPN, different sets of primers were designed to amplify genomic sequences of the gene. Only one SNP was identified in intron 4 (GenBank accession number NW_255516) using the primers OPNF: GCAAATCAGAAGTGTGATAGAC and OPNR: CCAAGCCAAACGTATGAT. Amplification of genomic DNA was performed in 25 μL of reaction volume, which included 50 ng of genomic DNA, 50 ng of each primer, 200 μM of each dNTP, 2.5 μL of 10x PCR buffer (Promega, Madison, WI), and 0.3 units of Taq DNA polymerase (Promega). The temperature cycles were as follows: 95°C for 5 min; 32 cycles of 94°C for 45 s, touchdown annealing from 63 to 50°C (−2°C/cycle) for 45 s, 72°C for 45 s; and a final extension at 72°C for 7 min. The PCR products were subjected to restriction using the enzyme BsrI that distinguishes alleles C and T of the SNP. The digestion products were electrophoresed on a 1.5% agarose gel; the T allele (uncut) was indicated by a band of 290 bp and the C allele was indicated by a band of 200 bp.

**Statistical Analyses**

For the CDDR data, maternal allele frequencies of OPN were estimated following Thaller et al. (2003), where all sons from homozygous sires and all homozygous sons of heterozygous sires were used. The maternal allele frequencies were estimated using the formula:

$$P_c = \frac{n_{CC} + n'_{TC}}{n_{CC} + n_{TC} + n'_{TT} + n_{TT}}$$

where $n_{CC}$ and $n_{TT}$ are the numbers of homozygous CC and TT sons within heterozygous sires; $n_{TC}$ and $n'_{TT}$ are the numbers of heterozygous TC and homozygous TT sons from homozygous TT sires. For the UW herd population, the allele frequencies were estimated by counting the number of each allele in the sample of 214 cows.

Weighted least squares analysis was employed to study the effects of OPN variants on production and functional traits in both the CDDR and UW herd populations. The model was

$$y_{ij} = \mu + Sire_i + \beta x_{ij} + e_{ij},$$

where $y_{ij}$ is the PTA of the trait that was considered for son (CDDR) or daughter (UW herd) $j$ of sire $i$, Sire$_i$
is the fixed effect of sire i, $\beta$ is the regression coefficient representing half of the allele substitution effect ($\alpha/2$); $x_{ij}$ is the number of C alleles (0, 1, or 2) for the jth son or daughter of sire i, and $e_{ij}$ is the residual. Reliabilities of the sons’ PTA were incorporated as weights in the model to obtain weighted least squares estimates for the allele substitution effects.

**RESULTS AND DISCUSSION**

Table 2 shows the distribution of genotypes of sons and cows for the CDDR and UW herd populations, respectively. For the CDDR population, the number of sons per grand sire family ranged from 19 to 102, with an average of 49 sons per family. Seven sires were homozygous CC; 7 sires were homozygous TT; and 14 sires were heterozygous. The estimated maternal C allele frequency was 0.52 (±0.02). The frequencies of C and T alleles in the UW herd were 0.49 and 0.51, respectively. Thus, the frequencies of OPN alleles seemed to be evenly distributed in both populations.

Estimated regression coefficients on the number of copies of the C allele (half of the allele substitution effects, $\alpha/2$) and their standard errors for production and health traits in the CDDR and UW herd populations are given in Table 3. For the CDDR population, the C allele was associated with an increase in milk protein percentage ($P = 0.0255$) and milk fat percentage ($P = 0.0480$). The correlation between the 2 traits was 0.57 in the CDDR population (Khatib et al., 2005). The OPN variants did not show significant effects on milk, fat, or protein yields or SCS. Although not statistically significant, allele C showed a negative effect on milk yield. This effect was expected because of the negative correlation ($-0.40$) between this trait and milk protein percentage.

For the UW herd population, the estimates of the effects of allele C were in the same direction (negative for milk yield and positive for milk protein percentage) as for the CDDR population, although these estimates did not reach a level of statistical significance. This could be due to the small number of animals (214) that was available for genotyping and phenotyping and low reliabilities of PTA for the cows (Table 1). However, the results of the UW herd did not contradict our findings in the CDDR population. An additional observation of note was that the C allele did not show any significant unfavorable effects on the other examined traits.

Our results are consistent with other studies that have shown a significant association of microsatellite markers in the region of OPN with milk protein percentage and other correlated traits (Zhang et al., 1998; Mosig et al., 2001; Nadesalingam et al., 2001; Ron et al., 2001; Rodriguez-Zas et al., 2002; Ashwell et al., 2004; Olsen et al., 2004). Recently, Olsen et al. (2005) positioned a QTL affecting milk production traits to an interval of 420 kb between the genes ABCG2 (ATP-binding cassette, subfamily G (WHITE), member 2) and LAP3 (leucine aminopeptidase 3) on bovine chromosome 6. This narrow region harbors only 6 genes, including OPN. While this study was being completed, Schnabel et al. (2005) reported that OPN was associated with milk protein percentage in the CDDR population. They searched for SNP in a region about 5 kb upstream of the bovine OPN and identified 6 SNP, of which 1 SNP (a deletion/insertion) showed significant association with milk protein percentage.

Although the causative mutation was likely not found in our study or in that of Schnabel et al. (2005), we conclude that either OPN itself affects milk protein percentage or it is in linkage disequilibrium with other

<table>
<thead>
<tr>
<th>Sire genotype</th>
<th>Son genotype</th>
<th>CC</th>
<th>CT</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td></td>
<td>136</td>
<td>181</td>
<td>0</td>
</tr>
<tr>
<td>CT</td>
<td></td>
<td>181</td>
<td>392</td>
<td>196</td>
</tr>
<tr>
<td>TT</td>
<td></td>
<td>0</td>
<td>156</td>
<td>120</td>
</tr>
</tbody>
</table>

Table 3. Estimates of the allele substitution effects and standard errors (SE) for production and health traits in the CDDR and UW herd populations.

<table>
<thead>
<tr>
<th>Trait</th>
<th>CDDR</th>
<th>UW herd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha/2$ (SE)</td>
<td>$P$</td>
</tr>
<tr>
<td>Milk yield</td>
<td>-28 (24)</td>
<td>0.2491</td>
</tr>
<tr>
<td>Milk fat yield</td>
<td>0.86 (0.88)</td>
<td>0.3229</td>
</tr>
<tr>
<td>Milk fat %</td>
<td>0.008 (0.004)</td>
<td>0.048</td>
</tr>
<tr>
<td>Milk protein yield</td>
<td>0.12 (0.60)</td>
<td>0.8481</td>
</tr>
<tr>
<td>Milk protein %</td>
<td>0.004 (0.002)</td>
<td>0.0255</td>
</tr>
<tr>
<td>SCS</td>
<td>-0.002 (0.006)</td>
<td>0.7165</td>
</tr>
</tbody>
</table>

1CDDR = Cooperative Dairy DNA Repository; UW = University of Wisconsin herd.

2Estimated regression coefficient $\beta$ representing half of the allele substitution effect ($\alpha/2$).
gene(s) that do. Further investigation of the OPN gene, including upstream and downstream control regions, is needed to elucidate molecular mechanisms causing the QTL effects.

ACKNOWLEDGMENTS
This research was supported by Hatch grant #WIS04736 from the University of Wisconsin. We thank the USDA Bovine Functional Genomics Laboratory staff for providing semen samples.

REFERENCES