Callipyge mutation affects gene expression in cis: A potential role for chromatin structure

Susan K. Murphy, Catherine M. Nolan, Zhiqing Huang, Katerina S. Kucera, Brad A. Freking, Timothy P.L. Smith, Kreg A. Leymaster, Jennifer R. Weidman and Randy L. Jirtle

Genome Res. 2006 16: 340-346; originally published online Jan 13, 2006;
Access the most recent version at doi:10.1101/gr.4389306

References
This article cites 26 articles, 8 of which can be accessed free at:
http://www.genome.org/cgi/content/full/16/3/340#References

Article cited in:
http://www.genome.org/cgi/content/full/16/3/340#otherarticles

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here

Notes

To subscribe to Genome Research go to:
http://www.genome.org/subscriptions/

© 2006 Cold Spring Harbor Laboratory Press
Callipyge mutation affects gene expression in cis:
A potential role for chromatin structure

Susan K. Murphy,1,2 Catherine M. Nolan,3 Zhiqing Huang,2 Katerina S. Kucera,1 Brad A. Freking,4 Timothy P.L. Smith,4 Kreg A. Leymaster,4 Jennifer R. Weidman,1 and Randy L. Jirtle1,5

1Department of Radiation Oncology, Duke University, Durham, North Carolina 27710, USA; 2Department of OB/GYN, Duke University, Durham, North Carolina 27708, USA; 3Zoology Department, University College Dublin, Belfield, Dublin 4, Ireland; 4U.S. Department of Agriculture, Agricultural Research Service, U.S. Meat Animal Research Center, Clay Center, Nebraska 68933-0166

Muscular hypertrophy in callipyge sheep results from a single nucleotide substitution located in the genomic interval between the imprinted Delta, Drosophila, Homolog-like 1 (DLK1) and Maternally Expressed Gene 3 (MEG3). The mechanism linking the mutation to muscle hypertrophy is unclear but involves DLK1 overexpression. The mutation is contained within CLPG1 transcripts produced from this region. Herein we show that CLPG1 is expressed prenatally in the hypertrophy-responsive longissimus dorsi muscle by all four possible genotypes, but postnatal expression is restricted to sheep carrying the mutation. Surprisingly, the mutation results in nonimprinted monoallelic transcription of CLPG1 from only the mutated allele in adult sheep, whereas it is expressed biallelically during prenatal development. We further demonstrate that local CpG methylation is altered by the presence of the mutation in longissimus dorsi of postnatal sheep. For 10 CpG sites flanking the mutation, methylation is similar prenatally across genotypes, but doubles postnatally in normal sheep. This normal postnatal increase in methylation is significantly repressed in sheep carrying one copy of the mutation, and repressed even further in sheep with two mutant alleles. The attenuation in methylation status in the callipyge sheep correlates with the onset of the phenotype, continued CLPG1 transcription, and high-level expression of DLK1. In contrast, normal sheep exhibit hypermethylation of this locus after birth and CLPG1 silencing, which coincides with DLK1 transcriptional repression. These data are consistent with the notion that the callipyge mutation inhibits perinatal nucleation of regional chromatin condensation resulting in continued elevated transcription of prenatal DLK1 levels in adult callipyge sheep. We propose a model incorporating these results that can also account for the enigmatic normal phenotype of homozygous mutant sheep.

Callipyge (from the Greek: calli-, beautiful; pyge, buttocks) is the name of a muscle hypertrophy syndrome that affects some descendents of a ram (named Solid Gold) born to a flock of sheep in Oklahoma in 1983. This ram produced several offspring that exhibited enlarged muscles resulting from hypertrophy of fast-twitch muscle fibers, accompanied by a substantial decrease in total body fat. Investigation of the inheritance pattern of the callipyge phenotype revealed that it is only evident when transmitted from the sire; inheritance from an affected ewe does not result in callipyge. This suggested that genomic imprinting is somehow involved, since imprinted genes are expressed from only one allele in a parent-of-origin-dependent manner. Indeed, linkage mapping of the callipyge allele showed it was located at the distal end of ovine chromosome 18 (Cockett et al. 1994) in a region syntenic to an imprinted domain in humans and mice. However, callipyge is unique in that homozygous mutant sheep do not exhibit the muscle hypertrophy phenotype, indicating a more complicated biology than can be explained by standard imprinting models. This phenotype represents the first known example of what was then coined “polar overdominance” (Cockett et al. 1996).

We recently identified the mutation, a single A-to-G base change located intergenic to the two imprinted genes DLK1 (also referred to as PREP-1, FAI, pG2) and MEG3 (also referred to as GTL2) that gives rise to the callipyge phenotype (Fig. 1; Freking et al. 2002). The causality of this mutation was subsequently supported by an independent study (Smit et al. 2003). DLK1 is a paternally expressed gene encoding a member of the epidermal growth factor repeat-like family and is related to the delta-notch family of proteins (Laborda et al. 1993; Laborda 2000). MEG3 is expressed from the maternally inherited chromosome and encodes multiple alternatively spliced transcripts that do not contain obvious open reading frames and are therefore thought to function as noncoding RNAs (Miyoshi et al. 2000).

We and others have determined that genes contained within this domain (see Fig. 1), including DLK1, MEG3, DLK1 associated transcript (DAT), Paternally expressed gene 11 (PEG11; also referred to as RTL1 in mice), and Maternally expressed gene 8 (MEG8), are imprinted in sheep (Charlier et al. 2001a; Murphy et al. 2005). While the presence of the callipyge mutation does not alter imprinting of these genes (Charlier et al. 2001a; Murphy et al. 2005), transcription levels of the genes linked in cis with the mutation are substantially influenced (Bidwell et al. 2001, 2004;
Callipyge mutation attenuates DNA methylation

Figure 1. Structure of the genomic region spanning the callipyge locus. Schematic representation of the telomeric end of ovine chromosome 18, and the genes present within the proximity of the callipyge mutation (A→G) on both the maternal (MAT) and paternal (PAT) chromosomes. Gene designations in mice are given in parentheses. The callipyge mutation (position 103,894 of AF354168) is contained within the transcribed CLPG1 locus. The direction of transcription is designated by the arrowheads representing each gene, with dark arrowheads representing actively transcribed alleles and translucent arrowheads representing silenced alleles for imprinted genes.

Results and Discussion

To begin to investigate how CLPG1 might be involved in the callipyge phenotype, we analyzed CLPG1 expression in tissues from four adult NMATCPAT (N and C refer to the normal and callipyge alleles, respectively, and the superscript designates the parental origin of the allele: MAT, maternally derived; PAT, paternally derived) sheep using RT-PCR (Fig. 2). We detected CLPG1 expression most consistently in muscles, including longissimus dorsi and biceps femoris, that are composed of fast-twitch fibers and undergo hypertrophy in the callipyge sheep.

Since imprinted genes flank CLPG1, we next determined whether CLPG1 is itself imprinted. We analyzed cDNAs from fetal and adult sheep heterozygous for the callipyge mutation (Fig. 3A; C, callipyge allele; N, normal allele; the maternal allele is listed first). In fetal sheep, CLPG1 was expressed from both parental alleles. When we analyzed the adult heterozygotes, we were surprised to find that CLPG1 is monoallelically expressed (Fig. 3B). Furthermore, CLPG1 is expressed from only the maternal allele in the NMATNPAT adults and conversely from only the paternal allele in the NNMATCPAT sheep. We also analyzed expression of the transcript in biceps femoris, another hypertrophy responsive muscle, of two NMATNPAT and three NMATCPAT adults, and found that the transcript was likewise expressed monoallelically from the maternal and paternal alleles, respectively (data not shown).

This unusual result prompted us to determine the expression status of this transcript in normal adult sheep. We identified several NMATNPAT sheep that were heterozygous for a G/A single nucleotide polymorphism (SNP) located 34 bp from the position of the callipyge mutation and also contained within the CLPG1 transcript. However, we were repeatedly unable to produce CLPG1 amplicons by RT-PCR from the longissimus dorsi muscle in these animals, while larger GAPDH amplicons were produced in each case (Fig. 4; data not shown). This was confirmed for a total of eight NMATNPAT animals and indicated that CLPG1 is normally not expressed in the NMATNPAT adults. When we examined CLPG1 expression in prenatal longissimus dorsi, we found that CLPG1 was, indeed, expressed in the NMATNPAT as well as the other three genotypes (Fig. 4). Although the homozygous NMATCPAT animals we examined were not informative at this second SNP, we did find that the NMATCPAT adults also express CLPG1 (Fig. 4). Together, these results suggest that CLPG1 is expressed during fetal development from both parental alleles, but expression is normally silenced postnatally in animals not carrying the callipyge mutation. Interestingly, the origin of nonimprinted monoallelic expression in the adult heterozygotes correlates with the presence of the mutation, suggesting that the mutation is somehow linked to CLPG1 transcription. This would predict that the NMATCPAT adult sheep express CLPG1 biallelically, but unfortunately because of the lack of polymorphisms in these inbred animals, it is not possible to confirm this at present.

One established mechanism to regulate allele-specific gene expression is via cytosine methylation, with generally a positive...
correlation between gene silencing and methylation. Therefore, we next examined whether methylation differences were present across the CLPG1 region in all four genotypes that might account for the observed differential expression. The area encompassing the callipyge mutation is not classified as a CpG island using standard criteria (observed/expected > 0.6; average percent GC > 0.5; >200 bp in length). Nevertheless, it does contain several CpG dinucleotides that might contribute to the regulation of this region (Fig. 5A). We used bisulfite sequencing and PhosphorImager analysis to quantify methylation of the CpG dinucleotides in the immediate vicinity of the mutation in two prenatal and two postnatal animals of each genotype (Fig. 5B,C). Figure 5B shows the average methylation detected for each CpG position, while the bar graphs in Figure 5C show the average methylation detected for all CpGs combined. During fetal development, there were no significant differences between the genotypes in longissimus dorsi when the average percent methylation for the entire region was calculated (ANOVA P = 0.9289). However, in normal adult sheep (N\textsuperscript{MAT}NP\textsuperscript{AT}), there was an approximate twofold increase in the level of methylation compared to that seen in the fetal N\textsuperscript{MAT}NP\textsuperscript{AT} sheep (63.2% vs. 32.1%, respectively), indicating peri- or postnatal acquisition of increased methylation for these CpG dinucleotides. Remarkably, we did not observe the same proportionate increases in methylation in the adult sheep carrying the callipyge mutation (Fig. 5B,C). Compared to the N\textsuperscript{MAT}NP\textsuperscript{AT} adults, the heterozygotes exhibited intermediate levels of methylation (C\textsuperscript{MAT}NP\textsuperscript{AT}, 48.48%; P = 0.05; N\textsuperscript{MAT}CP\textsuperscript{AT}, 46.59%, P = 0.04), and the homozygous adult C\textsuperscript{MAT}CP\textsuperscript{AT} sheep have the lowest methylation level (39.88%,
Callipyge mutation attenuates DNA methylation

We therefore examined the methylation status of these same sheep. RNA from each individual was reverse-transcribed in the presence or absence (no RT) of reverse transcriptase using a gene-specific primer, followed by PCR amplification and resolution by agarose gel electrophoresis. The expected amplicons are 270 bp and are designated by the arrowheads. GAPDH amplicons (452 bp) were amplified from the same tissues. (Lane M) 50-bp marker; (lane 1) N\text{MATCPAT}; (lane 2) C\text{MATCPAT}; (lane 3) N\text{MATNPAT}; (lane 4) C\text{MATNPAT}; (lane W) water control; (lane E) empty lane.

Figure 4. Prenatal and postnatal expression of CLPG1 in sheep of all four genotypes. RNA from each individual was reverse-transcribed in the presence or absence (no RT) of reverse transcriptase using a gene-specific primer, followed by PCR amplification and resolution by agarose gel electrophoresis. The expected amplicons are 270 bp and are designated by the arrowheads. GAPDH amplicons (452 bp) were amplified from the same tissues. (Lane M) 50-bp marker; (lane 1) N\text{MATNPAT}; (lane 2) C\text{MATNPAT}; (lane 3) N\text{MATCPAT}; (lane 4) C\text{MATCPAT}; (lane W) water control; (lane E) empty lane.

\[ P = 0.001 \text{ vs. } N\text{MATNPAT} \text{ adults; ANOVA } P = 0.01 \]. These striking results indicate a postnatal hindrance in the ability to acquire de novo methylation across this genomic locus in the presence of the mutation.

Because callipyge is specific to fast-twitch muscles, we hypothesized that the observed methylation patterns would only be observed in affected tissues if they have relevance to callipyge. We therefore examined the methylation status of these same sites in liver tissues from prenatal and postnatal animals of all four genotypes (Fig. 5B,C). The methylation level is much higher overall in liver than in longissimus dorsi, with normal prenatal sheep exhibiting an average 72.5% methylation, while postnatal sheep exhibit an average 76.3% methylation. Unlike longissimus dorsi, there is not a perinatal doubling of the methylation levels across this locus in liver, and there are not significant differences in methylation between the genotypes of adult sheep (ANOVA \( P = 0.19 \)). These results indicate that the attenuation of methylation acquisition observed in longissimus dorsi for this locus is specific to the affected tissues and results from the presence of the mutation.

One confounding aspect of the underlying mechanism of callipyge is the apparently normal phenotype of the homozygous C\text{MATNPAT} sheep. Georges et al. (2003) proposed that \textit{trans} interactions between the genes or gene products within the callipyge region would account for the incomplete penetrance of the phenotype in these animals. Our results support this contention. As shown schematically in Figure 6, we believe that MEG3 may act in \textit{trans} to down-regulate DLK1 expression, since we previously determined that the expression of DLK1 relative to MEG3 is elevated in the affected adult animals and is highly correlated with the callipyge phenotype (Murphy et al. 2005). This is also supported by the finding that mice inheriting a deletion of the imprint control region upstream of \textit{Gtl2} (MEG3) on the maternal allele exhibit loss of imprinting for all examined genes in the domain. In this case, \textit{Gtl2} expression is up-regulated and originates from both parental alleles, suggesting that transcription of \textit{Gtl2} or other maternally expressed genes is required to regulate expression of \textit{Dlk1} (Lin et al. 2003). Others have also postulated a role for \textit{PEG11} (Rhl1) and RNA interference originating from transcripts of the \textit{PEG11-AS} (Rhl1-AS) locus in the callipyge phenotype (Bidwell et al. 2004; Davis et al. 2005), although increasing evidence points toward \textit{Dlk1} as the effector of callipyge (Davis et al. 2004; Kim et al. 2004; Murphy et al. 2005).

Based on our results here, we refine this model by incorporation of the concept that chromatin structure, altered by the presence of the callipyge mutation, is integrally involved in mediating the callipyge phenotype. We have shown that in the N\text{MATNPAT} adult sheep, the intergenic DNA at the CLPG1 locus exhibits elevated postnatal cytosine methylation, presumably reflecting the results of chromatin remodeling. This is accompanied by CLPG1 silencing and a reduction in DLK1 mRNA levels (Murphy et al. 2005). In the N\text{MATCPAT} animals, the presence of the mutation on the paternally inherited chromosome prevents effective chromatin remodeling, and we have shown that CLPG1 and DLK1 are expressed at abnormally elevated levels in adulthood. In this case, MEG3 transcripts produced from the nonmutated maternally inherited chromosome are insufficient to abrogate DLK1 expression in \textit{trans}, resulting in the callipyge phenotype. For the C\text{MATCPAT} sheep, expression of MEG3 is enhanced, as would also be expected for DLK1 because of the homozygous mutational status. In these animals, however, the elevated levels of MEG3 transcripts control DLK1 expression in \textit{trans}, and thus these sheep do not have the callipyge phenotype. Our results further suggest that this regulation occurs at the transcriptional level, since the abundance of DLK1 mRNAs in the C\text{MATCPAT} sheep is equivalent to that observed in N\text{MATCPAT} sheep (Murphy et al. 2005).

The monoallelic transcription of CLPG1 only from the mutated allele in the adults suggests that the mutation is somehow allowing for constitutive regional gene expression from the affected allele. This is supported by the observation that while both DLK1 and MEG3 remain imprinted, transcription in adult animals is substantially increased from the allele carrying the mutation (Murphy et al. 2005). This long-range effect is consistent with the notion that the mutation may affect enhancer/silencer function, as previously proposed (Georges et al. 2003). Alternatively, it may affect chromatin structure with regional consequences, as we hypothesize here (Fig. 6). Increased postnatal methylation concomitant with decreased DLK1 and CLPG1 transcription in normal adult sheep suggests that chromatin structure is an important regulatory mechanism in the control of gene expression for this domain. Our results indicate that the callipyge mutation causes postnatal hypomethylation of this locus and prohibits down-regulation of CLPG1 transcription. This single base change also causes enhanced transcription of genes linked in \textit{cis} to the mutation but does not cause defects in imprinting. We therefore hypothesize that the callipyge mutation modifies
binding of a fast-twitch muscle-specific protein or protein complex that serves to nucleate regional chromatin condensation. As a result, elevated transcription is observed for genes within reach of this effect including \( DLK1 \), \( MEG3 \), and \( CLPG1 \). In support of this hypothesis, Charlier (2004) recently reported detection of a protein in sheep that binds to the region in which the callipyge mutation is located, and that binding of this protein is reduced when the mutation is present. Further work will be required to determine the identity of this protein and its role in the callipyge phenotype.

At this point, it would appear that the \( CLPG1 \) transcript itself does not have a causal role in the callipyge phenotype, since it is expressed in unaffected infraspinatus of the callipyge adult sheep. It is likely that the gene products produced from this locus are not functionally significant but rather are simply the result of open chromatin and the associated permissive conditions for transcription. Indeed, we have also detected an as-yet-uncharacterized transcript produced across this same region but from the opposite coding strand using gene-specific primers for reverse transcription (data not shown). Our data at this point are consistent with the mutation causing regional deregulation of gene expression through induced chromatin alterations.

These data have helped illuminate some of the enigmatic aspects of the callipyge phenotype and have contributed to our understanding of how a seemingly innocuous SNP has the potential to induce a considerable phenotypic change. Future work will be required to discern whether and how \( MEG3 \) (or other gene products within the callipyge domain) regulates \( DLK1 \) in trans and the specific nature of the hypothesized chromatin alterations induced by the presence of the mutation. Whether callipyge can be modeled (e.g., in mice engineered to carry the mutation) to further address questions about the biological basis for the effects of the callipyge mutation remains to be determined.

**Methods**

**Tissues**

Fetal and adult sheep tissues from all four callipyge genotypes were obtained from flocks maintained at the United States Department of Agriculture Meat Animal Research Center in Clay Center, Nebraska. Matt Doumit (Department of Animal Science, Michigan State University, East Lansing, MI) kindly provided additional muscle tissues from adult CMATNPAT animals from the same flocks, for confirmation of the allelic expression of \( CLPG1 \).

**Nucleic acid purification**

Genomic DNA was prepared from frozen tissues by proteinase K digestion (Qiagen) followed by standard phenol:chloroform extraction with Phase Lock Heavy gel separation tubes (Eppendorf) and ethanol precipitation. Total RNA was isolated from frozen tissues using RNA Stat-60 according to the manufacturer's directions (TelTest).

**Methylation analysis**

Genomic DNA was modified by conversion of unmethylated cytosines to uracils using sodium bisulfite as previously described (Grunau et al. 2001; Murphy et al. 2003). Bisulfite-treated DNA was subjected to PCR amplification using primers BS-F1 (5'-GGTGATTAGTAGTAGTAGTAGAGATA-3') and BS-R1 (5'-
Callipyge mutation attenuates DNA methylation

CTATTAATCCTAAATAATCCCTCT-3′) (94°C for 3 min, 40 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec, then 5 min at 72°C). The amplicons were resolved on agarose gels and purified using Sigma GenElute spin columns (Sigma Aldrich).

To determine the overall level of methylation, nucleotide sequence analysis was performed using the Radiolabeled Terminator Cycle Sequencing Kit (USB Corporation) with primer BS-F2 (5′-GTTTGGTCCTAA-3′), followed by polyacrylamide gel electrophoresis and determination of the percent cytosine methylation—(methylated C)/[(unmethylated C) + (methylated C)]—by PhosphorImager analysis (Molecular Dynamics Storm with ImageQuant software; Amersham Biosciences). A G/A polymorphism is present at one of the CpG dinucleotides within the region amplified by the primers (position 103,860 of Accession no. AF354168), and the recent generation of the callipyge allele occurred on a haplotype containing the A allele from this nearby single nucleotide polymorphism. The presence of the A at this position prevents methylation of the adjacent cytosine, and therefore analysis of the methylation status of this dinucleotide was not included in determining the methylation level of this region (Fig. 5C). Statistical comparisons across groups were performed using One-way Analysis of Variance (ANOVA) and the Dunnett Multiple Comparisons Tests (InStat; GraphPad Software). Comparisons between genotypes were performed using unpaired Student’s t-tests with a two-tailed distribution. P-values <0.05 were considered significant.

**CLPG1 expression analysis**

Using the callipyge mutation (position 103,894 of AF354168) and an upstream SNP as markers (position 103,860 of AF354168), heterozygous sheep were analyzed for allelic expression of CLPG1 by comparisons of the nucleotide sequence of genomic DNA to matched cDNA. Total RNA was treated with DNase I (Invitrogen) and subjected to reverse transcription in the presence and absence of reverse transcriptase (SuperScript II; Invitrogen) using a strand-specific primer (5′-CCAGAGGCTGC CAGGAGACC-3′), followed by PCR using forward primer F2 (5′-CAGATGCT GTGGACGCTGGAATC-3′) and reverse primer R244 (5′-GCAAGGTCT CTGTTGCTCCTAAA-3′) with conditions as follows: 94°C for 3 min; then 45 cycles of 94°C for 30 sec, 66°C for 30 sec, 72°C for 30 sec; and a 5-min extension at 72°C. GAPDH was used as an internal control and was amplified from oligo(dT)-primed cDNA generated from the same RNA preparations as those used for the CLPG1 analysis. Forward primer GAPDH-452F (5′-ACCAGAGTC CATGCCCACAC-3′) and reverse primer GAPDH-452R (5′-TCCACCCCTGTT GTGCTGA-3′) were used to generate 452-bp amplicons from cDNA as follows: 94°C for 3 min; then 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec; then a 5-min extension at 72°C. The 270-bp CLPG1 products were resolved by agarose gel electrophoresis, purified using Sigma GenElute columns (Sigma-Aldrich), then sequenced using the Radiolabeled Terminator Cycle Sequencing Kit (USB Corporation) with primer R3 (5′-GAT CACCCCAACCCCTGGC-3′). Sequencing products were resolved on 5% polyacrylamide sequencing gels and exposed to radiographic film and/or to a PhosphorImager screen for analysis using ImageQuant software (Amersham Biosciences).

**Acknowledgments**

We thank Matt Dowmoot from the Department of Animal Science, Michigan State University, East Lansing, Michigan for kindly providing additional C*M*N*P*F*H*1 adult sheep tissues for confirmation of the expression status of CLPG1 transcripts. We also thank Ali Gusbarg at Duke University (Durham, NC), and Renee Godtel and Kevin Tennill at the MARC location (Clay Center, NE) for excellent technical support contributions. This research was supported by NIH grants CA94668 to S.K.M., Enterprise Ireland grant IC/2003/20 to C.M.N., and EA25951, ES08823, and ES13053 to R.L.J.

**References**

Cockett, N., Georges, M., and Charlier, C. 2004. Expression of DLK1 protein in skeletal muscle of padumnal phenotype, the only known example of polar overdominance in mammals.