Toward Molecular Understanding of Polar Overdominance at the Ovine Callipyge Locus

M. Georges,* C. Charlier,† M. Shm,‡ E. Davis,‡ T. Shay,§ X. Torrode,∥ H. Takeda,* F. Caiment,* and N. Cockett†

*Department of Genetics, Faculty of Veterinary Medicine, University of Liege (B43), 4000 Liege, Belgium;
†Department of Animal, Dairy and Veterinary Sciences, College of Agriculture, Utah State University, Logan, Utah 84322-4700

The callipyge phenotype (Gk. calli- = beautiful + -pyge = buttoks) is a generalized muscular hypertrophy described in sheep. It is due to an increase in the size and proportion of fast twitch muscle fibers. It manifests itself only after birth at ~1 month of age. It exhibits a rostro-caudal gradient being more pronounced in the muscle of the pelvic limb and torso, hence its name. It is accompanied by a decrease in all measures of fatness. Affected animals are characterized by an improved feed efficiency and dressing percentage (for review, see Cockett et al. 2001). Quite logically, the callipyge phenotype initially caught the attention of animal breeders because of its potential agronomic value. Ensuing studies, however, would reveal some remarkable features of the callipyge phenotype, especially its non-Mendelian mode of inheritance. These would quickly attract more attention among the scientific community than its potential economic value, especially since the quality of callipyge meat appeared to be mediocre!

The aim of this paper is to update and complement a recent review describing the present understanding of the genetics and epigenetics of the callipyge phenomenon (Georges et al. 2003).

Polar Overdominance at the Ovine CLPG Locus

The callipyge phenotype was first reported in the 1980s, showing in ~15% of offspring of a Dorset ram called “Solid Gold.” When mated to wild-type ewes, callipyge rams descending from Solid Gold produced 50% callipyge offspring, irrespective of sex. This Mendelian segregation ratio suggested that the callipyge phenotype resulted from an autosomal, nonrecessive mutation referred to as “CLPG” (Cockett et al. 1994). This monogenic hypothesis was confirmed when the CLPG locus was mapped to a 4.5-cM marker interval on distal chromosome 18 (Cockett et al. 1994; Freking et al. 1998; Shay et al. 2001).

Unexpectedly, crosses involving callipyge ewes and rams, each known to be of +/CLPG homozygous genotype, would not yield 50% callipyge offspring as expected in case of parental imprinting, but rather the unusual 75% wild-type versus 25% callipyge phenotypic ratio. Marker analysis indicated that only the +/CLPG heterozygous genotype was associated with the CLPG phenotype, the CLPG+/+ offspring being wild type although carrying the CLPG mutation on their paternal chromosome. This non-Mendelian inheritance pattern was referred to as “polar overdominance” (Cockett et al. 1996). It was postulated to result either from a mutation that would switch the imprinting of the CLPG gene from paternal to maternal expression, or would simultaneously knock out an imprintated maternally expressed trans-acting repressor and its target CLPG gene. In the former case, the callipyge phenotype would result from the illegitimate absence of the CLPG gene product, in the latter from its illegitimate presence. The latter model turns out to share many features with reality as we understand it today.

The CLPG Mutation Maps to the DLK1-GTL2 Imprinted Domain

A BAC contig spanning the CLPG locus was constructed (Segers et al. 2000; Berghmans et al. 2001; Shay et al. 2001), and ~500 contiguous kilobases predicted to contain the mutation were sequenced (Charlier et al. 2001b and unpubl.). In silico annotation of this sequence showed that the CLPG mutation mapped to the newly described DLK1-GTL2-imprinted domain (Fig. 1).

This evolutionary conserved domain, which spans ~1 Mb, harbors at least four protein-encoding genes with preferential expression from the paternal allele (BEGIN, ...
served exon–intron organization, extensive alternative splicing. The gene trap screen was used to identify several tissue-specific transcripts by means of alternative splicing. The most widely expressed, producing multiple transcripts—as a result of alternative promoter usage and splicing—that exhibit paternal or biallelic expression in a tissue- and promoter-specific manner (M.A. Smit et al., in prep.).

**BEGAIN** (brain-enriched guanylate kinase-associated protein) encodes a protein that binds to the guanylate kinase domain of PSD-95/SAP90, a scaffolding protein at the postsynaptic cell membrane (Deguchi et al. 1998). It is widely expressed, producing multiple transcripts—as a result of alternative promoter usage and splicing—that exhibit paternal or biallelic expression in a tissue- and promoter-specific manner (M.A. Smit et al., in prep.).

**DLK1** is a member of the **EGF** domain containing Notch/Delta/Serrate protein family whose function remains poorly understood. It has been implicated in adipogenesis, hematopoiesis, lymphopoiesis, and neuroendocrine differentiation, as well as tumorigenesis (Labenda et al. 2000). DLK1 null mice display accelerated adiposity, as well as symptoms shared with mUPD12 mice and mUPD14 humans: growth retardation, biphaphthalmosisis, and skeletal abnormalities (Moon et al. 2002). A possible involvement of DLK1 in myogenesis is suggested by the myofiber hypertrophy and delayed maturation observed in murine pUPD12 fetuses expressing a double dose of **DLK1** (Georgiadis et al. 2000), and by the known involvement of Notch signaling in myogenesis (see, e.g., Hirsinger et al. 2001; Conboy et al. 2003). **PEG11** corresponds to a long (~1300 residues), uninterrupted open reading frame (ORF) that has the potential to code for a protein with a central portion that is highly similar to the gag and pol central portion that is highly similar to the gag and pol domain in a range of tissues and throughout developmental stages was imprinting of any of the studied genes (**DLK1**, **PEG11**, **BEGAIN**, **GTL2**, **antiPEG11**, **MEG8**, and **MIRG**) affected by the **CLPG** mutation? We monitored the expression of the genes in the **DLK1-GTL2** domain in a range of tissues and throughout development for sheep representing the four possible **CLPG** genotypes. In none of the examined tissues and developmental stages was imprinting of any of the studied genes (**DLK1**, **PEG11**, **BEGAIN**, **GTL2**, **antiPEG11**, **MEG8**, and **MIRG**) affected by the **CLPG** mutation: mRNAs were always preferentially expressed from the paternal allele, while ncRNA genes were always preferentially expressed from the maternal allele (Charlier et al. 2001a; M.A. Smit et al., in prep. and unpubl.). Thus, this allowed us to exclude the first model for polycystic overdominance assuming a switch in imprinting of the **CLPG** gene from paternal to maternal expression. However, the expression levels in skeletal muscle of a cluster of centrally positioned genes (**DLK1**, **PEG11**, **GTL2**, **antiPEG11**, **MEG8**, and **MIRG**), but not **BEGAIN**
POLAR OVERDOMINANCE AT THE OVINE CALLIPYGE LOCUS

Figure 1. Schematic representation of the imprinted DLK1-GTL2 domain. The genes shown in red are protein-encoding genes that are preferentially transcribed from the paternal allele (Pat) in the direction indicated by the arrows. The genes shown in blue are non-coding RNA genes that are preferentially transcribed from the maternal allele (Mat) in the direction shown by the arrow. The C/D snoRNAs and mRNAs hosted, respectively, by MEG8, antiPEG11, and MIRG are shown underneath. The IG-DMR is an imprinted control element that is methylated (open circle) and inactive on the paternal allele, while unmethylated (closed circle) and active on the maternal chromosome, thereby leading to the expression of the ncRNA genes and silencing of the protein-encoding genes.

and DIO3) were clearly influenced by the CLPG genotype (Charlier et al. 2001a; M.A. Smit et al., in prep. and unpubl.). More specifically, the concentrations of mRNA (DLK1, PEG11) were markedly increased in the two genotypes sharing the CLPG mutation on their maternally inherited chromosome (CLPGMat/+Pat), while the concentrations of ncRNA (GTL2, antiPEG11, MEG8, and MIRG) were markedly increased in the two genotypes sharing the CLPG mutation on their paternally inherited chromosome (CLPGPat/+Mat) that were homozygous for all known polymorphisms within and in the vicinity of the CLPG locus. Three hundred and twenty polymorphisms were identified differentiating the resequenced + and − alleles from the reference BAC sequence. The resequenced alleles, on the contrary, were identical over their entire length with the exception of a single A (+ allele) to G (CLPG allele) transition located between the DLK1 and GTL2 genes, at 32.8 Kb from the closest genes, i.e., GTL2 (Freking et al. 2002; Smit et al. 2003).

Was this the CLPG mutation? A number of arguments strongly support this conjecture. First, the A to G transition affects the third base pair of a dodecamer motif that is perfectly conserved in the 13 mammalian species in which it has been sequenced. The extensive conservation of this motif strongly suggests that it fulfills an important function. The dodecamer motif is itself embedded in a 2-Kb segment exhibiting >70% similarity between human and sheep. Second, the G allele was only encountered in the callipyge flock, i.e., descendents of Solid Gold, despite the screening of more than 270 animals representing 13 distinct breeds, including wild-type Dorset. Third, and most convincingly, Solid Gold was shown to be mosaic being germ-line mosaic as well. This strongly suggests that the A to G transition occurred during Solid Gold’s early development and virtually proves that it is the causative mutation. To provide final proof of its causal-ity, we are generating transgenic mice with the corresponding nucleotide substitution by gene targeting and hope to recapitulate the callipyge phenotype and mode of inheritance in this more tractable model.

How might this putative cis-acting element operate? So far, classical approaches for the functional analysis of regulatory elements have provided relatively little insight. An ~500-bp fragment encompassing the wild-type or mutant dodecamer motif in its center does not seem to affect the expression level of a luciferase reporter gene driven either by the CMV or IGF2 P3 promoter in C2C12

Figure 1. Schematic representation of the imprinted DLK1-GTL2 domain. The genes shown in red are protein-encoding genes that are preferentially transcribed from the paternal allele (Pat) in the direction indicated by the arrows. The genes shown in blue are non-coding RNA genes that are preferentially transcribed from the maternal allele (Mat) in the direction shown by the arrow. The C/D snoRNAs and mRNAs hosted, respectively, by MEG8, antiPEG11, and MIRG are shown underneath. The IG-DMR is an imprinted control element that is methylated (open circle) and inactive on the paternal allele, while unmethylated (closed circle) and active on the maternal chromosome, thereby leading to the expression of the ncRNA genes and silencing of the protein-encoding genes.
ECTOPIC EXPRESSION OF DLK1 PROTEIN IN SKELETAL MUSCLE OF \( +^{\text{Mat}}/+^{\text{Pat}} \) INDIVIDUALS CAUSES THE CALLIPGYE PHENOTYPE

The observed effect of the CLPG mutation on transcript levels of neighboring genes in skeletal muscle did not satisfactorily explain why only \( +^{\text{Mat}}/+^{\text{Pat}} \) individuals express the callipygous phenotype. Indeed, \( -\text{CLPG}^{\text{Mat}}/+^{\text{Pat}} \) individuals share the overexpression of the \( \text{DLK1} \) and \( \text{PEG11} \) transcripts with \( \text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Mat}} \) individuals and the lack of overexpression of \( \text{GTL2} \), \( \text{PEG11} \), \( \text{MGEA} \), and \( \text{MIRG} \) with \( -\text{CLPG}^{\text{Mat}}/-\text{CLPG}^{\text{Pat}} \) individuals, neither of which exhibits callipygous features. It is thus the combination of mRNA overexpression and lack of ncRNA overexpression that seems unique.

To gain a better understanding of why this might be, we monitored the expression of \( \text{DLK1} \) at the protein level by immunohistochemistry (Davis et al. 2004). \( \text{DLK1} \) protein could not be detected in skeletal muscle of \( -\text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \), \( -\text{CLPG}^{\text{Mat}}/-\text{CLPG}^{\text{Mat}} \), and \( \text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Mat}} \) individuals, irrespective of muscle group and developmental stage. Remarkably, in \( +\text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \) individuals, \( \text{DLK1} \) protein was found abundantly, albeit exclusively, in skeletal muscle exhibiting the muscular hypertrophy, e.g., Longissimus dorsi after 1 month of age. At earlier developmental stages or in nonhypertrophied muscle groups, \( \text{DLK1} \) protein could not be detected in \( +\text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \) individuals either. We thus observed a perfect association between the expression of \( \text{DLK1} \) protein in skeletal muscle and their hypertrophy, both when differentiating individuals by \( \text{CLPG} \) genotype and when distinguishing muscle groups within \( +\text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \) individuals.

To test whether this association might be causal, we generated transgenic mice expressing the membrane-bound form of the \( \text{DLK1} \) protein (as we demonstrated that the corresponding \( \text{DLK1} \) transcripts are by far the most abundant in skeletal muscle of callipygous animals) in skeletal muscle under the dependence of a myosin light chain promoter and enhancer (Davis et al. 2004). The two transgenic lines that were produced indeed exhibited a muscular hypertrophy as a result of an increase in myofiber diameter. This strongly suggests that the ectopic expression of \( \text{DLK1} \) protein that is observed in hypertrophied muscle of callipygous sheep causes this hypertrophy or at least contributes to it. We are presently using the same transgenic based approach to test the putative effect on muscle mass of ectopic expression of \( \text{PEG11} \).

POLAR OVERDOMINANCE SUPPORTS THE TRANS-INTERACTION BETWEEN THE PRODUCTS OF RECIPROCALLY IMPRINTED GENES

The question remains why \( \text{DLK1} \) protein is detected in skeletal muscle of \( -\text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \) individuals but not in that of \( \text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Pat}} \) animals, while \( \text{DLK1} \) mRNA is present at comparable concentrations in skeletal muscle of both genotypes, transcribed from the paternal \( \text{CLPG} \) allele. Obviously the difference between the two genotypes is the maternal allele that is wild type and essentially silent in skeletal muscle of \( -\text{CLPG}^{\text{Mat}}/-\text{CLPG}^{\text{Pat}} \) individuals while being \( \text{CLPG} \) and thus producing ncRNA in skeletal muscle of \( \text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Mat}} \) animals. This thus suggests that in \( \text{CLPG}^{\text{Mat}}/\text{CLPG}^{\text{Mat}} \) individuals, the ncRNAs are blocking the translation of the \( \text{DLK1} \) mRNAs in trans explaining the absence of \( \text{DLK1} \) protein (Fig. 2).

It is particularly interesting in this regard that the \( \text{DLK1}-\text{GTL2} \) locus is remarkably rich in miRNA genes (cf. above). Could it be that one or several of these maternally expressed miRNAs target \( \text{DLK1} \) transcripts, thus mediating the translational trans inhibition postulated to underlie polar overdominance? To test this hypothesis we compared the affinity of 43 miRNAs predicted by Mirscan (Lim et al. 2003) in the \( \text{DLK1}-\text{GTL2} \) domain us- ing the human, murine, and rat sequence for the corresponding 3'UTR of \( \text{DLK1} \) (X. Tordoir et al., unpubl.). Targetscan was used to quantify the affinity of the miRNAs for their target (Lewis et al. 2003), and affinities were summed across species. As the boundaries of the miRNAs cannot be determined unambiguously, multiple candidate miRNAs were actually tested for each pri- miRNA. The affinity of the predicted miRNAs for the 3'UTR of \( \text{DLK1} \) was compared with their affinity for a set of 676 size-matched control 3'UTRs. The affinity of the miRNAs was either tested as a group (combinatorial rheostat hypothesis, Bartel 2004) or individually (individual rheostat hypothesis). We found no evidence sup- porting the fact that, as a group, the miRNAs predicted in the \( \text{DLK1}-\text{GTL2} \) have a higher affinity for the 3'UTR of \( \text{DLK1} \) than for the 3'UTR of a random set of control
When tested individually, however, we found one miRNA that had an affinity for DLK1 superior to 90% of the best affinities found with the miRNA set for the 3’UTR of the control genes. Noteworthy, this miRNA also obtained the highest Mirscan score (19). The potential interaction between this miRNA and DLK1 transcripts is being tested experimentally.

It is noteworthy that there is strong evidence for a trans interaction between the maternally expressed miR127 and miR136 miRNA genes embedded in antiPEG11 and the perfectly complementary paternally PEG11 transcripts.

First, Lin et al. (2003) provided indirect evidence by showing that concentrations of PEG11 transcripts were sixfold rather than twofold higher in mice inheriting the ∆IG-DMR deletion on their maternal chromosome when compared to wild-type controls. This is postulated to be due to the absence of miR127 and miR136 expression in the former. More directly, we have recently cloned PEG11 cleavage products from skeletal muscle of CLPGMat/CLPGPat sheep predicted to result from the action of miR127 (E. Davis et al., in prep.). This observation is in perfect agreement with the lower levels of PEG11 transcripts observed in CLPGMat/CLPGPat when compared to +/+CLPGMat animals. 

Although miRNAs are the most attractive candidate mediators of the postulated trans effect, alternative hypotheses should not be overlooked. It is intriguing in this regard that noncoding H19 transcripts might bind to the same IMP (IGF2 mRNA binding protein) postulated to regulate translatability of some IGF2 mRNAs (Runge et al. 2000). Might this point to a translational trans regulation of IGF2 mediated by H19 ncRNA? Might similar mechanisms operate at the DLK1-GTL2 domain?

CONCLUSIONS

The study of the callipyge phenomenon has provided some unique opportunities to probe the novel epigenetic mechanisms that underlie its unusual mode of inheritance—polar overdominance. More than being just an ovine idiosyncrasy, the study of polar overdominance is likely to shed light on molecular mechanisms that might be involved in the inheritance of other phenotypes, including complex inherited disorders in the human. The study of the callipyge phenotype once again illustrates the potential value of domestic animal biodiversity in unraveling fundamental biological processes (Andersson and Georges 2004).

ACKNOWLEDGMENTS

This work was supported by grants from (i) the FRFC (no. 2.4525.96), (ii) Crédit aux Chercheurs (no. 1.5.134.00) from the FNRS, (iii) Crédit à la Recherche from the U.L.g. (iv) the STCC (no. 0135), (v) the European Union (Callimat), (vi) the Utah Center of Excellence Program, (vii) the USDA/NRI/GRP (Grants #94-40435, #96-35205, and #98-03455), and (viii) the Utah Agricultural...
REFERENCES


