TOWARDS THE MOLECULAR UNDERSTANDING OF THE POLAR OVERDOMINANCE PHENOMENON ASSOCIATED WITH THE CALLIPYGE PHENOTYPE IN SHEEP

Par M^me le Dr. C. CHARLIER

Lauréate du prix pharma.be (période 2002-2003)

Introduction

The callipyge phenotype (CLPG) is a generalized muscular hypertrophy described in sheep. It is due to an increase in diameter and proportion of fast-twitch muscle fibers (Type II), is characterized by a rostro-caudal gradient and has a post-natal expression. Its expression is under the dependence of the callipyge locus and mutation, localized at the telomeric end of ovine chromosome 18 (Cockett *et al.,* 1994). The uniqueness of the callipyge phenotype results from its unusual inheritance pattern referred to as polar overdominance: only heterozygous individuals receiving the CLPG mutation from their sire (+^{MAT}/CLPG^{PAT}) exhibit the phenotype (Cockett *et al.,* 1996).

The CLPG locus was first mapped to a 400 Kb chromosome segment (*Berghmans *et al.,* 2001) contained within a BAC contig that we constructed for this region (*Shay *et al.,* 2001; *Segers *et al.,* 2000). It contains four protein-encoding genes, expressed preferentially from the paternal allele (DLK1, PEG11, BEGAIN2 and DIO3) as well as multiple non-coding RNA genes expressed preferentially from the maternal allele (GTL2, MEG8, multiple C/D snoRNA and miRNA) (*Charlier *et al.,* 2001a; *Paulsen *et al.,* 2001; Cavaille *et al.,* 2002; Seitz *et al.,* 2003).

Resequencing 200kb allowed us to identify the CLPG mutation as an A-to-G transition in a highly conserved dodecamer motif located 35Kb 5' from GTL2, the closest gene (Freking *et al.,* 2002; *Smit *et al.,* 2003).

The CLPG mutation was shown to inactivate a silencer element controlling -the expression of all the genes in the domain in cis, in skeletal muscle after birth, without altering their imprinting status. As a consequence, (+^{MAT}/CLPG^{PAT}) callipyge individuals are characterized by a unique expression profile: an overexpression of the mRNA coding for DLK1 and PEG11 in the absence of an overexpression of the non-coding RNAs (GTL2, MEG8, C/D snoRNA et miRNA) (*Charlier *et al.,* 2001b).

Despite equivalent levels of DLK1 mRNA in both +^{MAT}/CLPG^{PAT} and CLPG^{MAT}/ CLPG^{PAT} individuals, DLK1 protein is only detectable in skeletal muscles of +^{MAT}/CLPG^{PAT} individuals, which are the only ones expressing the callipyge phenotype. This demonstrates the existence in CLPG^{MAT}/CLPG^{PAT} individuals, of a post-transcriptional trans inhibition of DLK1 hypothetically mediated by one or more of the non-coding RNA genes expressed exclusively from the maternal CLPG allele (*Georges *et al.,* 2003; *Davis *et al.,* submitted).

To verify the hypothesis of a direct role of ectopic DLK1 expression in the determinism of the callipyge phenotype, we produced transgenic mice constitutively expressing DLK1 in skeletal muscle. These were shown to present a muscular hypertrophy reminiscent of the callipyge phenotype in sheep, allowing us to conclude that DLK1 is indeed the likely effector (*Davis *et al.,* submitted).

It is noteworthy that the callipyge phenotype has been elected «mutant of the month» by Nature Genetics (October 2003) and has been the subject of a recent review article on epigenetics in Scientific American (December 2003) (Figure 1).

The references marked by an * in this introduction are part of the presented work.

This manuscript, comprising five chapters supported by nine publications, recapitulates the positional cloning, identification and characterization of the cis and trans effects of the CLPG point mutation causing the callipyge phenotype in sheep.

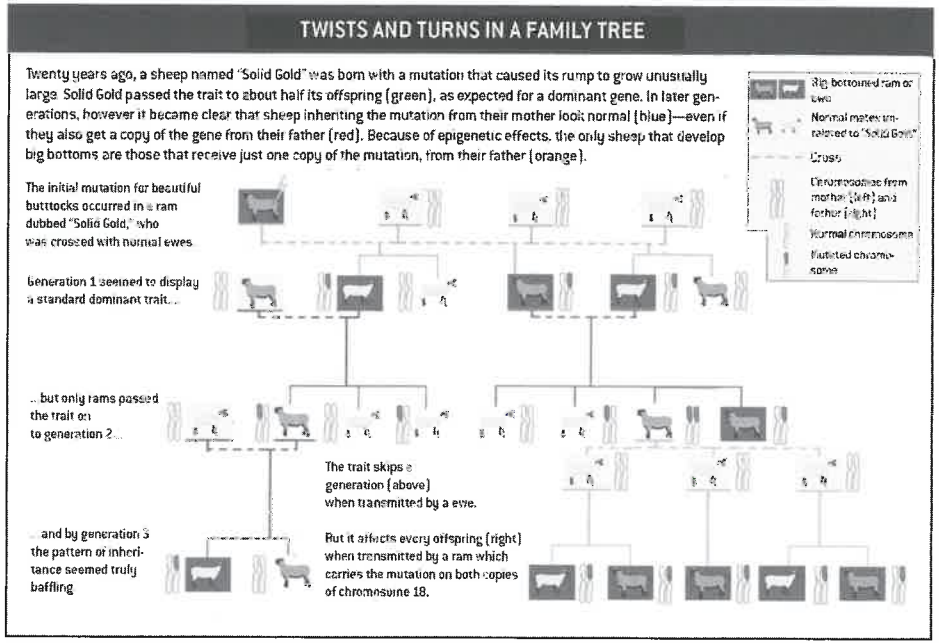

Figure 1: Polar overdominance phenomenon.
From Scientic American, December 2003

CHAPITRE I
Physical mapping and fine mapping of the callipyge mutation

SEGERS K., VAIMAN D., BERGHMANS S., SHAY T., BEEVER J., COCKETT N., GEORGES M. & CHARLIER C., *Construction and characterization of an ovine BAC contig spanning the callipyge locus*, Animal Genetics 31, 352-359 (2000).

We describe the construction of an ovine BAC contig spanning a 4.6 centimorgan (cM) chromosome segment known to contain the callipyge (CLPG) locus. The contig comprises 21 ovine BAC clones jointly covering approximately 900 kilobases (Kb). Two gaps in the BAC contig, spanning 10 and 7.5 Kb, respectively, were bridged by long range PCR. The corresponding chromosome region was shown to be characterized by an unusually low Kb to cM ratio (164 Kb/cM) and a high density of Not1 sites (1:126 Kb) possibly reflecting a high gene density in the corresponding chromosome region. Equivalent amplification of 64 sequence tagged sites spanning the corresponding region from homozygous +/+ and CLPG/CLPG individuals disproves the hypothesis of a major deletion causing the CLPG mutation.

SHAY T., BERGHMANS S., SEGERS K., MEYERS S., WOMACK J., BEEVER J., GEORGES M., CHARLIER C. & COCKETT N.E., *Fine-mapping and construction of a bovine contig spanning a 4.6 centimorgan interval containing the CLPG locus*, Mammalian Genome, 12, 141-149 (2001).

The callipyge (CLPG) gene was fine-mapped by linkage analysis to a 4.6-cM chromosome interval on distal ovine OAR18q, flanked by microsatellite markers IDVGA30 and OY3. The OAR18q linkage map and human HSA14q transcript map were aligned by genotyping two bovine-hamster whole-genome radiation hybrid panels with the microsatellite markers, as well as with sequences corresponding to HSA 14q genes. Using Type I loci mapping to the IDVGA30-OY3 interval as anchor points, we have constructed a 1.4-Mb bovine BAC contig containing the IDVGA30-OY3 interval. We demonstrate that the IDVGA30-OY3 interval spans approximately 770 kb and contains at least four genes: YY1, WARS, DLK1, and GTL2.

BERGHMANS S., SEGERS K., SHAY T., GEORGES M., COCKETT N.E. & CHARLIER C., *Breakpoint mapping positions the callipyge gene within a 400 kilobase chromosome segment containing the Dlk1 and Gtl-2 genes*, Mammalian Genome, 12, 183-185 (2001).

CHAPTER II
Comparative analysis of a new imprinted domain encompassing the callipyge mutation

CHARLIER C., SEGERS K., WAGENAAR D., KARIM L., BERGHMANS S., JAILLON O., SHAY T., WEISSENBACH J., COCKETT N., GYAPAY G., GEORGES M., *Human – ovine comparative sequencing of a 250 kilobase imprinted domain encompassing the callipyge (clpg) gene and identification of six imprinted transcripts: DLK1, DAT, GTL2, PEG11, antiPEG11 and MEG8*, Genome Research, 11, 850-862 (2001a).

Two ovine BAC clones and a connecting long-range PCR product, jointly spanning approximately 250 kb and representing most of the MULGE5-OY3 marker interval known to contain the clpg locus, were completely sequenced. The resulting genomic sequence was aligned with its human ortholog and extensively annotated. Six transcripts, four of which were novel, were predicted to originate from within the analyzed region and their existence confirmed experimentally: DLK1, DAT, GTL2, PEG11, antiPEG11, and MEG8. RT-PCR experiments performed on a range of tissues sampled from an 8-wk-old animal demonstrated the preferential expression of all six transcripts in skeletal muscle, which suggests that they are under control of common regulatory elements. The six transcripts were also shown to be subject to parental imprinting: DLK1, DAT, and PEG11 were shown to be paternally expressed and GTL2, antiPEG11, and MEG8 to be maternally expressed.

PAULSEN M., TAKADA S., YOUNGSON N.A., BENCHAIB M., CHARLIER C., SEGERS K., GEORGES M., FERGUSON-SMITH A.C., *Detailed sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species identifies highly conserved genomic elements and a domain structure different from the Igf2-H19 region*, Genome Research, 11, 2085-2094 (2001).

The Dlk1-Gtl2 domain on mouse chromosome 12 contains reciprocally imprinted genes with the potential to contribute to our understanding of common features involved in imprinting control. We have sequenced this conserved region in the mouse and sheep and included the human sequence in a three species comparison. This analysis resulted in a precise conservation map and identification of highly conserved sequence elements, some of which we have shown previously to be differentially methylated in the mouse. Additionally, this analysis facilitated identification of a CpG-rich tandem repeat array located approximately 13-15 kb upstream of Gtl2. Furthermore, we have identified a third imprinted transcript that overlaps with the last Dlk1 exon in the mouse. This transcript lacks a conserved open reading frame and is probably generated by cleavage of extended Dlk1 transcripts. Because Dlk1 and Gtl2 share many of the imprinting properties of the well-characterized Igf2-H19 domain, it has been proposed that the two regions may be regulated in the same way. Comparative genomic examination of the two domains indicates that although there are similarities, other features are very different, including the location of conserved CTCF-binding sites, and the level of conservation at regulatory regions.

CHAPTER III
Cis effects of the callipyge mutation

CHARLIER C., SEGERS K., KARIM L., SHAY T., GYAPAY G., COCKETT N.E., GEORGES M., *The callipyge (CLPG) mutation enhances the expression of the coregulated DLK1, GTL2, PEG11 and MEG8 genes in cis without affecting their imprinting status*, Nature Genetics, 27, 367-369 (2001b).

The callipyge (CLPG) phenotype described in sheep is an inherited muscular hypertrophy that is subject to an unusual parent-of-origin effect referred to as polar overdominance: only heterozygous individuals having inherited the CLPG mutation from their sire exhibit the muscular hypertrophy. The callipyge (clpg) locus was mapped to a chromosome segment of approximately 400 kb (refs. 2-4), which was shown to contain four genes (DLK1, GTL2, PEG11 and MEG8) that are preferentially expressed in skeletal muscle and subject to parental imprinting in this tissue. Here we describe the effect of

the CLPG mutation on the expression of these four genes, and demonstrate that callipyge individuals have a unique expression profile that may account for the observed polar overdominance.

CHAPTER IV
Identification of the callipyge mutation

SMIT M., SEGERS K., SHAY T., BARALDI F., GYAPAY G., SNOWDER G., GEORGES M., COCKETT N., CHARLIER C., *Mosaicism of Solid Gold supports the causality of a non-coding A to G transition in the determinism of the callipyge phenotype*, Genetics, 163, 453-456, (2003).

To identify the callipyge mutation, we have resequenced 184 kb spanning the DLK1-, GTL2-, PEG11-, and MEG8-imprinted domain and have identified an A-to-G transition in a highly conserved dodecamer motif between DLK1 and GTL2. This was the only difference found between the callipyge (CLPG) allele and a phylogenetically closely related wild-type allele. We report that this SNP is in perfect association with the callipyge genotype. The demonstration that Solid Gold — the alleged founder ram of the callipyge flock — is mosaic for this SNP, virtually proves the causality of this SNP in the determinism of the callipyge phenotype.

CHAPTER V
Trans effects at the callipyge locus

GEORGES M., CHARLIER C., COCKETT N., *Polar overdominance at the ovine callipyge locus supports trans interaction between the products of reciprocally imprinted genes*, Trends in Genetics, 19, 248-252, (2003).

The callipyge phenotype in sheep is an inherited muscular hypertrophy that affects only heterozygous individuals who receive the CLPG mutation from their father. The CLPG mutation is a single nucleotide substitution in what is probably a long-range control element (LRCE) within the DLK1-GTL2 imprinted domain. Recent results suggest that the unique mode of inheritance of callipyge, referred to as polar overdominance, results from the combination of the cis-effect of the CLPG mutation on the expression levels of genes in the DLK1-GTL2 imprinted domain, and the trans interaction between the products of reciprocally imprinted genes.

DAVIS E., HARKEN JENSEN C., DAA SCHRÖDER H., SHAY T., KLIEM A., COCKETT N., GEORGES M., CHARLIER C., *Ectopic expression of DLK1 protein in skeletal muscle of padumnal heterozygotes causes the callipyge phenotype*, Submitted for publication, (2004).

The callipyge phenotype is an inherited skeletal muscle hypertrophy described in sheep. It is characterized by an unusual mode of inheritance («polar overdominance») in which only heterozygous individuals having received the CLPG mutation from their father express the phenotype 1. The CLPG mutation was shown to be an A to G transition

in a highly conserved dodecamer motif located between the imprinted DLK1 and GTL2 genes 2,3. This motif is thought to be part of a long range control element (LRCE) as the CLPG mutation was shown, in skeletal muscle, to enhance the transcript levels of the DLK1, PEG11, GTL2 and MEG8 genes in cis without altering their imprinting status 4. As a result the $+^{MAT}/CLPG^{PAT}$ individuals have a unique expression profile assumed to underlie the callipyge phenotype: an overexpression of the protein encoding DLK1 and PEG11 transcripts, in the absence of an overexpression of the non-coding GTL2 and MEG8 transcripts. However, the way in which this distinct expression profile causes the callipyge muscular hypertrophy remained unclear.

RÉSUMÉ

VERS LA COMPRÉHENSION MOLÉCULAIRE DU PHÉNOMÈNE DE SURDOMINANCE POLAIRE ASSOCIÉ AU PHÉNOTYPE CALLIPYGE DU MOUTON

De nombreux phénotypes d'intérêt médical et agronomique sont de nature héréditaire dite «complexe»: bien qu'influencés par des gènes, ils ne se transmettent pas selon les lois «simples» de Mendel. Récemment Steve Henikoff écrivait: «The nature of quantitative-trait variation is one of the last unexplored frontiers in genetics, awaiting the future and definitive identification of complex trait determinants, wether they be genetic or epigenetic» (Nature Genetics, 2003).

Le phénotype callipyge est une hypertrophie musculaire du mouton caractérisée par un mode de transmission «complexe» qualifié de surdominance polaire: seuls les individus hétérozygotes ayant hérité la mutation CLPG de leur père expriment le phénotype. Au cours des dix dernières années, nous nous sommes consacrés à élucider les bases moléculaires de ce phénomène unique. Ces travaux ont abouti à un modèle comprenant les éléments suivants: (i) DLK1, un promoteur de croissance exprimé exclusivement à partir de l'allèle paternel, (ii) un inhibiteur en trans de DLK1 exprimé exclusivement à partir de l'allèle maternel et correspondant vraisemblablement à un ARN non-codant, (iii) la mutation CLPG invalidant un élément «silenceur» réprimant l'expression du promoteur et de son inhibiteur en trans dans le muscle squelettique.

Outre l'intérêt fondamental de ces recherches, la dissection moléculaire de la surdominance polaire contribuera à une meilleure compréhension et donc à un meilleur contrôle des pathologies «complexes» caractérisées par des effets dits «parent of origin».

SUMMARY

TOWARDS THE MOLECULAR UNDERSTANDING OF THE POLAR OVERDOMINANCE PHENOMENON ASSOCIATED WITH THE CALLIPYGE PHENOTYPE IN SHEEP

Most medically or agronomically important phenotypes are «complex» inherited traits. They are influenced by genes but not transmitted according to Mendel's laws. Recently, Steve Henikoff wrote: «The nature of quantitative-trait variation is one of the last unexplored frontiers in genetics, awaiting the future and definitive identification of complex trait determinants, wether they be genetic or epigenetic» (Nature Genetics, 2003).

The callipyge phenotype is a muscular hypertrophy in sheep that is characterized by a complex inheritance pattern referred to as polar overdominance: only heterozygous individuals having received the CLPG mutation from their sire exhibit the phenotype. We have spent the last ten years attempting to dissect the molecular basis of this unique phenomenon. These studies have lead to a working model including: (i) DLK1, a paternally expressed growth promoter, (ii) a maternally expressed DLK1 trans-acting repressor

that is likely to be a non-coding RNA, and (iii) the CLPG mutation inactivating a silencer element controlling both the expression of DLK1 and its trans-acting repressor in skeletal muscle.

This research is not only of fundamental interest, but may as well lead to a better understanding and hence control of «complex» diseases exhibiting «parent-of-origin» effects.

BIBLIOGRAPHIE

1. CAVAILLE J., SEITZ H., PAULSEN M., FERGUSON-SMITH A.C., *Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region*, Human Molecular Genetics, 13,1527-1538 (2002).
2. COCKETT N., JACKSON S., SHAW T., FARNIR F., SNOWDER G., NIELSEN D., GEORGES M., *Polar overdominance at the ovine callipyge locus*, Science, 273, 236-238 (1996).
3. COCKETT N.E., JACKSON S.P., SHAY T.D., NIELSEN D., GREEN R.D., GEORGES M., *Chromosomal localization of the callipyge gene in sheep (Ovis aries) using bovine DNA markers*, Proceedings of the National Academy of Sciences, US, 91, 3019-3023 (1994).
4. FREKING B.A., MURPHY S.K., WYLIE A.A., RHODES S.J., KEELE J.W., LEYMASTER K.A., JIRTLE R.L., SMITH T.P., *Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals*, Genome Research, 12, 1496-1506 (2002).
5. SEITZ H., YOUNGSON N., LIN S.P., DALBERT S., PAULSEN M., BACHELLERIE J.P., FERGUSON-SMITH A.C., CAVAILLE J., *Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene*, Nature Genetics, 34,261-262 (2003).

(Faculté de Medecine Vétérinaire, Département des Productions Animales, U.Lg).

*
* *