

The callipyge mutation enhances the expression of coregulated imprinted genes in *cis* without affecting their imprinting status

The callipyge (CLPG) phenotype (from $\kappa\alpha\lambda\iota$, "beautiful," and $\pi\iota\gamma\epsilon$, "buttocks") 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 hypertrophy¹. 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^{5–9}. 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.

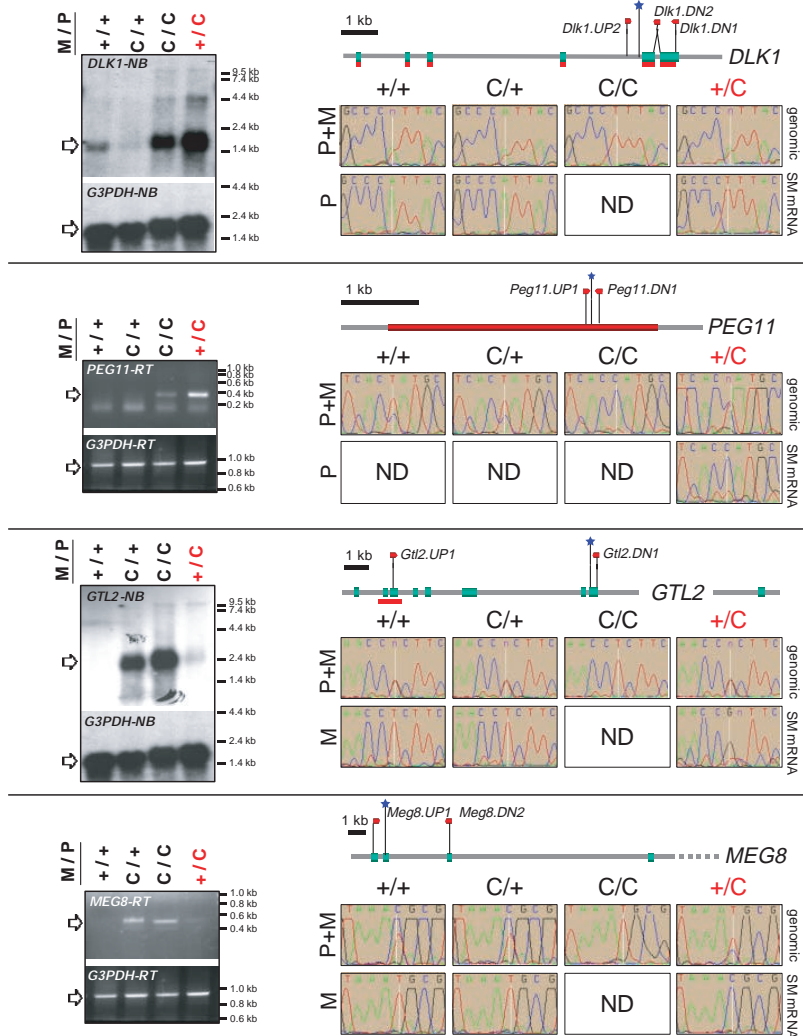
Using the available marker and phenotypic information, we identified individuals representing the four possible genotypes at the *clpg* locus from our callipyge flock²: $+/+$, $CLPG^M/+^P$, $+^M/CLPG^P$ and $CLPG/CLPG$ (where the 'M' and 'P' superscripts refer to the maternal and paternal homologs, respectively). We selected animals on the basis of their heterozygosity for previously described single-nucleotide polymorphisms (SNP) in the transcribed portions of the studied

genes⁵, and our ability to unambiguously trace the parental origin of the respective alleles. This was possible for individuals of

all genotypes, except $CLPG/CLPG$, which were homozygous by descent for all SNPs at the *clpg* locus because all known CLPG chromosomes descend from a common founder ram². The selected animals were slaughtered at eight weeks, an age at which they fully express the callipyge phenotype. We prepared RNA from longissimus dorsi, a skeletal muscle that is hypertrophied in callipyge individuals, and used it to examine the expression of the imprinted genes *DLK1*, *GTL2*, *PEG11* and *MEG8* in the four possible *clpg* genotypes.

We repeatedly showed by northern blot and/or RT-PCR that transcripts from the two genes which are paternally expressed and encode a protein product (*DLK1*, encoding an EGF-like homeotic protein family member involved in adipogenesis, and *PEG11*, encoding a gag and pol-like polyprotein of unknown function) are very abundant in skeletal muscle of $CLPG/CLPG$ and $+^M/CLPG^P$ individuals (that is, the two genotypes that have the CLPG mutation on their paternal

Fig. 1 Evaluation of the transcript levels and imprinting status of *DLK1*, *PEG11*, *GTL2* and *MEG8* in skeletal muscle (longissimus dorsi) of eight-week individuals representing the four possible callipyge genotypes: wild type, $CLPG^M/+^P$, $+^M/CLPG^P$ and $CLPG/CLPG$. Transcript levels were estimated by northern-blot analysis or RT-PCR using standard procedures and primer combinations⁵. The presence of similar amounts of RNA for the four callipyge genotypes was demonstrated by ethidium bromide staining of the gels (data not shown), hybridization with a *G3PDH* probe, or RT-PCR with *G3PDH*-specific primers (*G3PDH.UP1*, 5'-TGAAGGTCGTTGAACG-GATTGGC-3'; *G3PDH.DN1*, 5'-CATGTAGGCCAT-GAGGTCCACCAC-3'). The imprinting status of the corresponding transcripts was studied by cycle-sequencing the corresponding RT-PCR products. Sequence traces obtained from genomic DNA (containing therefore the paternal (P) and maternal (M) alleles) and skeletal muscle mRNA (containing either P or M) are shown. White vertical lines within the electropherograms mark the position of the used SNPs (ref. 5). The RT-PCR amplicons from $CLPG/CLPG$ individuals were not sequenced as these are not informative (ND, not done). The structure of the corresponding genes is schematized, including the positions of the used SNPs (blue star), primers used for RT-PCR and/or cycle-sequencing (small red arrows) and sequences used as probes on northern blots (red lines underneath the gene). Exonic sequences are shown as green boxes (*DLK1*, *GTL2*, *MEG8*) when transcribed in the cent to ter direction, or as red boxes (*PEG11*) when transcribed in the ter to cent direction. See ref. 5 for more details about the organization of the callipyge imprinted domain.



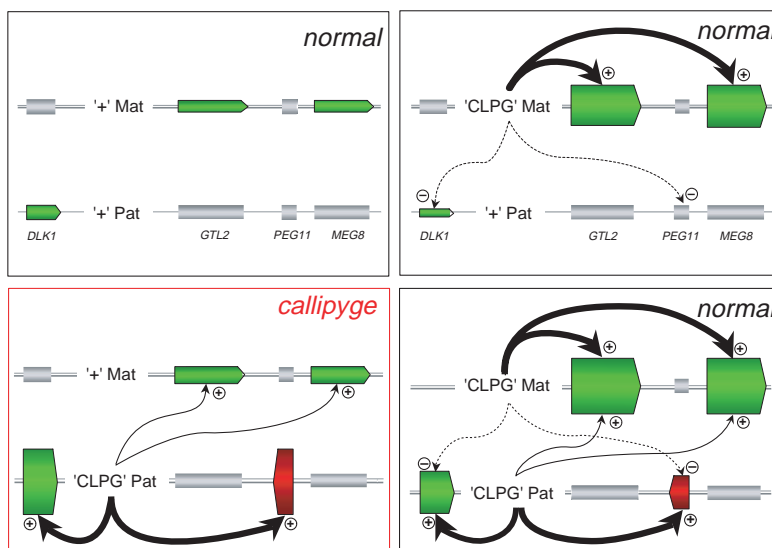


Fig. 2 A model for the callipyge polar overdominance. The expression profile of *DLK1*, *GTL2*, *PEG11* and *MEG8* are illustrated for $+/+$, $CLPG^{M/+P}$, $+^M/CLPG^P$ and $CLPG/CLPG$ individuals. The $+^M/CLPG^P$ individuals expressing the callipyge phenotype are boxed in red. Genes transcribed in the cen to ter direction are represented as green arrows, those transcribed in the ter to cen direction as red arrows. The thickness of the arrows reflects the transcription level. The identified *cis* effects are illustrated by the thick black arrows connecting the hypothetical location of the CLPG mutation and the influenced genes. The putative *trans* effects are likewise illustrated by the thin black arrows.

chromosome in common) compared with skeletal muscle of $+/+$ and $CLPG^{M/+P}$ individuals, in which they were either present at very low concentration (*DLK1*) or essentially absent (*PEG11*; Fig. 1a,b). We studied the imprinting status of the corresponding mRNAs by cycle-sequencing RT-PCR products encompassing previously described *DLK1* and *PEG11* SNPs (ref. 5) for which the selected individuals were heterozygous. Cycle sequencing the *DLK1* amplicons obtained from $+/+$, $CLPG^{M/+P}$ and $+^M/CLPG^P$ individuals demonstrated the presence of the paternal *DLK1* allele only, irrespective of the *clpg* genotype of the examined individual (Fig. 1a). Cycle sequencing the *PEG11* amplicon obtained from $+^M/CLPG^P$ individuals likewise indicated the exclusive presence of the paternal allele (no *PEG11* RT-PCR product was obtained for $+/+$ and $CLPG^{M/+P}$ individuals; Fig. 1b). These results indicate that the CLPG mutation enhances the expression of *DLK1* and *PEG11* in *cis* without altering their imprinting status. Note that similar results were obtained for the previously described transcript *DAT* (for *DLK1* associated transcript), which was also shown to be imprinted and expressed exclusively from the paternal allele⁵ (data not shown).

Using the same approach, we showed that transcripts from the two genes which are maternally expressed and encode an RNA product (*GTL2* and *MEG8*) are very

abundant in $CLPG^{M/+P}$ and $CLPG/CLPG$ individuals (that is, the two genotypes that have the CLPG mutation on their maternal chromosome in common) compared with wild-type and $+^M/CLPG^P$ individuals (Fig. 1c,d). As for *DLK1* and *PEG11*, we studied the imprinting status of the corresponding mRNAs by cycle sequencing RT-PCR products encompassing previously described *GTL2* and *MEG8* SNPs (ref. 5) for which the selected individuals were heterozygous. Cycle sequencing the *GTL2* and *MEG8* amplicons obtained from $+/+$, $CLPG^{M/+P}$ and $+^M/CLPG^P$ individuals clearly demonstrated the presence of the maternal allele only, irrespective of the *clpg* genotype of the examined individual (Fig. 1c,d). Therefore, these results indicate that the CLPG mutation enhances the expression of *GTL2* and *MEG8* in *cis* without altering their imprinting status. Preliminary analyses indicate a similar effect of the CLPG mutation on the previously described imprinted anti-*PEG11* transcripts with maternal expression⁵ (data not shown).

Our results show that the CLPG mutation does not alter the imprinting status of *DLK1*, *PEG11*, *GTL2* and *MEG8* in the callipyge domain, but enhances their expression level in *cis*. This indicates that the CLPG mutation modifies the activity of a common regulatory element, which could be either an enhancer (boosted by

the CLPG mutation) or a silencer (inhibited by the CLPG mutation).

The only genotype to express the callipyge phenotype (that is, $+^M/CLPG^P$ individuals) differs from the three other *clpg* genotypes by the overexpression of *DLK1* and *PEG11* in skeletal muscle, in the absence of a high expression of *GTL2* and *MEG8*. We hypothesize that this unique expression profile is underlying the observed polar overdominance (Fig. 2). The callipyge muscular hypertrophy would result from an increase in *DLK1* and/or *PEG11* activity in skeletal muscle due to the effect of the paternally inherited CLPG mutation on their respective transcription levels. In $CLPG/CLPG$ individuals, concomitant overexpression of *GTL2* and *MEG8* could interfere in *trans* with *DLK1* and *PEG11* activity and thereby block phenotypic expression. This interference could operate at the transcriptional level (competition for transcription factors), on the stability of the *DLK1/PEG11* mRNAs, on their translatability or even at the protein level.

Closer examination of the transcript levels of the studied genes indeed provides preliminary evidence for such *trans* effects (Figs. 1 and 2). When maternally inherited, the CLPG mutation decreases the expression level of the paternally expressed genes. As a consequence, *DLK1* and *PEG11* transcript levels are superior in $+^M/CLPG^P$ versus $CLPG/CLPG$ individuals. On the other hand, the paternally inherited CLPG mutation seems to increase the expression level of the maternally expressed genes. As the same tendency was repeatedly observed on independent samples (data not shown), we do not believe that it would only reflect trivial sampling variation, but rather think that it corresponds to a real effect of *clpg* genotype.

It is unclear at this point how an overexpression of *DLK1* and/or *PEG11* might cause the observed muscular hypertrophy. The observation that they are preferentially expressed in skeletal muscle⁵, however, strengthens this hypothesis. Overexpressing the respective genes in skeletal muscle using transgenic techniques should help to unravel the precise contribution of each of these genes to the callipyge phenotype.

Acknowledgments

We thank A. Ferguson-Smith and M. Paulsen for discussions. This project was supported by grants from Fonds de la Recherche Fondamentale Collective (n° 2.4525.96), Crédit aux Chercheurs (n° 1.5.134.00) from the Fonds National de la Recherche Scientifique, Crédit à la Recherche from the University of Liège, the Utah Center of



Excellence Program, the USDA/NRI Competitive Grants Program (USDA/NRICGP grants 94-04358, 96-35205 and 98-03455) and the Utah Agricultural Experiment Station.

Carole Charlier^{1*}, Karin Segers^{1*}, Latifa Karim¹, Tracy Shay³, Gabor Gyapay², Noelle Cockett³ & Michel Georges¹

*These authors contributed equally to this work.

¹Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), Liège, Belgium. ²Genoscope, Centre National de Séquençage, Evry, France. ³Department of Animal, Dairy and Veterinary Sciences, College of Agriculture, Utah State University, Logan, Utah, USA. Correspondence should be addressed to M.G. (e-mail: michel.georges@ulg.ac.be).

Received 30 November 2000; accepted 2 March 2001.

1. Cockett, N.E. *et al. Science* **273**, 236–238 (1996).
2. Shay, T. *et al. Mamm. Genome* **12**, 141–149 (2001).
3. Segers, K. *et al. Anim. Genet.* **31**, 352–359 (2001).
4. Berghmans, S. *et al. Mamm. Genome* **12**, 183–185 (2001).
5. Charlier, C. *et al. Genome Res.* (in press).
6. Miyoshi, N. *et al. Genes Cells* **5**, 211–220 (2000).
7. Schmidt, J.V. *et al. Genes Dev.* **14**, 1997–2002 (2000).
8. Takada, S. *et al. Curr. Biol.* **10**, 1135–1138 (2000).
9. Wylie, A.A. *et al. Genome Res.* **10**, 1711–1718 (2000).

Mutations in *SIP1*, encoding Smad interacting protein-1, cause a form of Hirschsprung disease

Hirschsprung disease (HSCR) is sometimes associated with a set of characteristics including mental retardation, microcephaly, and distinct facial features^{1–3}, but the gene mutated in this condition has not yet been identified. Here we report that mutations in *SIP1*, encoding Smad interacting protein-1, cause disease in a series of cases. *SIP1* is located in the deleted segment at 2q22 from a patient with a *de novo* t(2;13)(q22;q22) translocation. *SIP1* seems to have crucial roles in normal embryonic neural and neural crest development.

Patients with HSCR associated with microcephaly, mental retardation, hypertelorism, submucous cleft palate and short stature (MIM 235730) show genetic heterogeneity. Some are sibling cases, suggestive of autosomal recessive disease^{1,3–7}, and others are isolated cases likely to have contiguous gene syndromes or dominant mutations^{2,7–10}. We have collected more than 200 cases of HSCR. Among them, five patients presented with HSCR associated with microcephaly, mental retardation, epilepsy and characteristic facial features. These five patients were all isolated cases, and three of them were complicated by congenital heart disease (patent ductus arteriosus and/or ventricular septal defect). They had birth weights between 2,820 and 3,200 grams. Patient 1 also had a t(2;13)(q22;q22) translocation (Fig. 1). Microcephaly, hypertelorism, convergent strabismus and wide nasal bridge were observed in all cases, with facial features similar to those previously reported^{2,8}. Deafness, pigmentation

defects, iris coloboma, ptosis and cleft palate were not observed.

Because there are many similarities in the clinical presentations between our patients and the *Ece1*-null mouse¹¹, as well as a patient with an Arg742Cys missense mutation in *ECE1* (ref. 12), we first investigated *ECE1* as the candidate gene. Detailed mutational analyses including Southern blot and direct nucleotide sequencing of all the exons and the flanking splice junctions failed to reveal any causative mutations in our HSCR patients. We next analyzed the *de novo* translocation breakpoint at 2q22 of patient 1 because two previously reported patients^{2,9} had interstitial deletions in chromosome 2 (del 2q21–23 and del 2q22–23), indicating that the causative gene may be located at this breakpoint (Fig. 1). To determine the precise location of the translocation breakpoint, we carried out fluorescence *in situ* hybridization (FISH) using several cDNAs and RP11-BAC clones located in

this area and loss of heterozygosity (LOH) studies using polymorphic microsatellite markers. The results showed patient 1 to have an approximately 5-Mb cytogenetic deletion flanked by *D2S129* and *D2S151* on 2q22 (Fig. 2). This deleted segment on 2q22 is more than 2 Mb distant from *CACNB4*, an epilepsy-associated gene, indicating that *CACNB4* is not involved in the etiology of epilepsy of our patient. The chromosomal breakpoint on 13q22 was also determined to be located within an approximately 1-Mb interval between *EDNRB*, a gene known to be mutated in HSCR, and *SHGC-14012* (BAC clone, 360P17), which is downstream of *EDNRB*, indicating lack of *EDNRB* involvement in the etiology.

The genes *KYNU* (encoding kynureninase), *PRO0159* and *SIP1*, have been mapped to this region (Fig. 2). We focused on *SIP1* as a candidate gene because overexpression of wild-type *SIP1* in *Xenopus laevis* results in abnormal head phenotypes¹³. *SIP1* spans approximately 70 kb and consists of 10 exons and 9 introns. Comparison of the human and mouse homologs of *SIP1* at nucleotide and amino acid levels revealed 93% and 97% similarities, respectively. We searched for mutations in *SIP1* in 4 patients (patients 2, 3, 4 and 5) with direct nucleotide sequencing analysis of PCR products obtained from the 10 exons and the flanking splice junctions. We found two nonsense mutations (1645A→T resulting in R549X, patient 2; 2083C→T resulting in R695X, patient 4) and a 4-bp deletion (1173del AACA resulting in 392fs394X, patient 5) in exon 8 (Fig. 2 and Web Fig. A). We did not find any mutations in this gene in patient 3. The possibility of mutations in the regulatory regions, introns or splice sites, however, cannot be completely excluded. The finding that the parents of patients 1, 2, 4 and 5 are all healthy with no mutations in *SIP1* indicates the identified mutations to be *de novo* events.

Sip1, isolated from the mouse embryo as a Smad1-interacting protein using the yeast two-hybrid system, is a member of

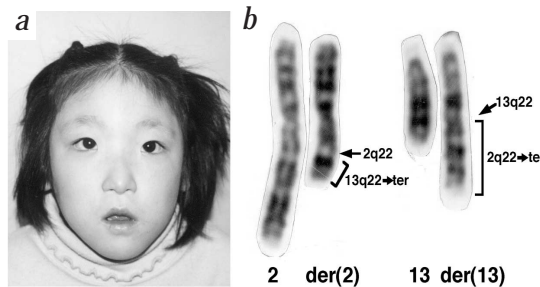


Fig. 1 Physical and cytogenetic examinations. **a**, Facial appearance of patient 1 at 5 years of age. **b**, Cytogenetic analysis of peripheral blood lymphocytes revealed a reciprocal translocation between chromosome 2 and chromosome 13: 46,XX,t(2;13)(q22;q22).