nature genetics

A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep

Alex Clop[1,6], Fabienne Marcq[1,6], Haruko Takeda[1,6], Dimitri Pirottin[1,6], Xavier Tordoir[1], Bernard Bibé[2], Jacques Bouix[2], Florian Caiment[1], Jean-Michel Elsen[2], Francis Eychenne[2], Catherine Larzul[2], Elisabeth Laville[3], Françoise Meish[1], Dragan Milenkovic[4], James Tobin[5], Carole Charlier[1] & Michel Georges[1]

Texel sheep are renowned for their exceptional meatiness. To identify the genes underlying this economically important feature, we performed a whole-genome scan in a Romanov × Texel F2 population. We mapped a quantitative trait locus with a major effect on muscle mass to chromosome 2 and subsequently fine-mapped it to a chromosome interval encompassing the myostatin (*GDF8*) gene. We herein demonstrate that the *GDF8* allele of Texel sheep is characterized by a G to A transition in the 3′ UTR that creates a target site for *mir1* and *mir206*, microRNAs (miRNAs) that are highly expressed in skeletal muscle. This causes translational inhibition of the myostatin gene and hence contributes to the muscular hypertrophy of Texel sheep. Analysis of SNP databases for humans and mice demonstrates that mutations creating or destroying putative miRNA target sites are abundant and might be important effectors of phenotypic variation.

To locate quantitative trait loci (QTL) underlying the muscular hypertrophy of Texels, we generated a Romanov × Texel F2 with 258 offspring. We chose a hypermuscled Belgian strain of Texel for these crosses. We examined 37 phenotypes measuring body composition, muscularity and fat deposition on the F2 animals[1,2]. We genotyped 153 microsatellites and scanned the genome by linear regression assuming one QTL per chromosome and fixation of alternate (*T* and *R*) QTL alleles in the parental breeds[3]. A QTL with major effect on muscularity was identified on chromosome 2 (OAR2). For traits yielding genome-wide $P < 0.05$, the QTL accounted for 5–25% of the variance, the difference between alternate homozygotes ($2a$) ranged from 0.68 to 1.66 residual standard deviations, and the dominance deviation ranged from $-0.70a$ to $0.50a$. The QTL explained one-fifth to one-third of the difference between parental breeds (**Supplementary Table 1** online). The confidence interval (c.i.) for the QTL spanned 10 cM (**Fig. 1a**). 'Within-family' analyses showed that the three F1 rams were heterozygous *TR* (**Supplementary Fig. 1** online). A QTL with similar effects and location has been detected in other Texel-based pedigrees[4].

To refine the map position of this QTL we generated a higher-density map of the c.i. In a first approach (marker-assisted segregation analysis (MASA)), we produced 43 offspring from an F2 ram (20254) that inherited an intact Texel chromosome and a chromosome recombining within the *DIK4864–BM81124* interval. The recombinant chromosome was of Romanov descent distal from *DIK4864* and of Texel descent proximal from *BM81124*. The 'weight of the hindquarters' was 238 g heavier in the offspring that inherited the Texel chromosome than in those that inherited the recombinant chromosome ($P = 0.0013$) (**Fig. 1b**). This difference was similar to the *R* to *T* substitution effect estimated in the F2 animals ($a = 203$ g). The effects on all other measured traits were very similar to those observed in the F2 animals (**Supplementary Fig. 2**). This indicated that the ram was *TR* and the QTL located distal from *DIK4864* (**Fig. 1b**).

In the second approach, we hypothesized that selection for meatiness in Belgian Texel animals might have caused near-fixation of a favorable QTL allele *T* ('selective sweep') and that most hypermuscled Texels would be homozygous for a segment encompassing the QTL. To test this, we genotyped 42 hypermuscled Texels as well as 108 controls (16 breeds) for ten microsatellites spanning the c.i. We measured the increase in frequency of a given marker allele in the Texels with respect to controls using DISLAMB[5]. The likelihood ratio test maximized on the OAR2q side of the c.i., in agreement with the MASA (**Fig. 1c**). The signal peaked at marker *BULGE20*: one of its 15 alleles had a frequency of 94% in Texels but only 12% in controls; heterozygosity was 11% in Texels versus 85% in controls.

[1]Unit of Animal Genomics, Department of Animal Production, Faculty of Veterinary Medicine & Centre for Biomedical Integrative Genoproteomics, University of Liège (B43), 20 Boulevard de Colonster, 4000 Liège, Belgium. [2]Institut National de la Recherche Agronomique–Station d'Amélioration Génétique des Animaux (INRA-SAGA), BP 52627, 31326 Castanet-Tolosan CEDEX, France. [3]Station de Recherches sur la Viande, INRA, Theix, 63122 Saint-genès-Champanelle, France. [4]INRA/Université de Limoges, Faculté des Sciences, 87060 Limoges Cedex, France. [5]Cardiovascular and Metabolic Diseases, Wyeth Research, 87 Cambridge Park Drive, Cambridge, Massachusetts 02140, USA. [6]These authors contributed equally to this work. Correspondence should be addressed to M.G. (michel.georges@ulg.ac.be).

Received 11 December 2005; accepted 26 April 2006; published online 4 June 2006; doi:10.1038/ng1810

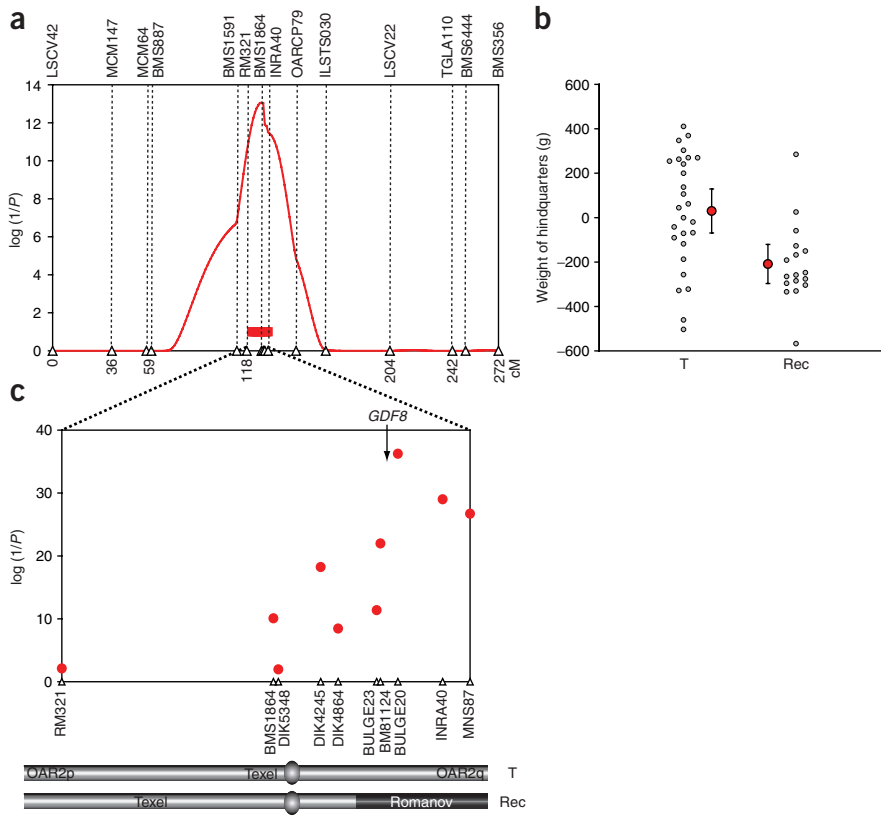

Figure 1 Mapping and fine-mapping of a QTL influencing muscularity on sheep chromosome 2. (**a**) QTL location scores (expressed as log(1/*P*), where *P* corresponds to the genome-wide *P* value of the data under the null hypothesis of no QTL) obtained for 'weight of the hindquarters (in g)' along the marker map of sheep chromosome 2 (OAR2). The red horizontal bar corresponds to the 95% c.i. for the QTL location obtained by bootstrapping. (**b**) Marker-assisted segregation analysis. Gray circles: 'weight of the hindquarters' (in g) (residuals after correction for fixed effects) of offspring, sorted according to the chromosome 2 inherited from their sire: 'T' (full-length Texel chromosome) or 'Rec' (recombinant Texel-Romanov chromosome; the position of the recombination breakpoint is shown in **c**). Red circles represent the group phenotypic mean (± 1.96 s.e.m. of the estimate). The higher variance in the T group is in agreement with the recessive effect of the T allele on weight of the hindquarters (see **Supplementary Table 1**). (**c**) Results of the DISLAMB[5] linkage disequilibrium analysis in the 95% c.i. for the QTL to identify the effects of a putative selective sweep. Log(1/*P*) corresponds to the logarithm (base 10) of 1/*P*, where *P* is the likelihood of the data under the null hypothesis of no enrichment of a specific marker allele in Texel when compared with controls. The position of the *GDF8* gene is shown by the arrow. The structures of 'T' and 'Rec' chromosomes of the ram used in **b** are schematically represented.

In cattle, the myostatin (*GDF8*) gene is located in the *BM81124–BULGE20* interval. *GDF8* loss-of-function mutations cause double-muscling in mice, cattle and humans[6,7], making it an obvious candidate. We sequenced the coding regions of the *GDF8* gene from DNA of three F0 Texel rams and seven controls (five F0 Romanov ewes, one Dorset and one Tarasconnais), but we did not find any polymorphism. RNA blots showed a band of the expected size with comparable intensity in both Texels and controls (**Fig. 2a**). Real-time quantitative RT-PCR did not demonstrate any significant breed effect on *GDF8* RNA levels (**Supplementary Fig. 3** online). We amplified the *GDF8* ORF by RT-PCR on RNA derived from muscle, sequenced the corresponding PCR products and confirmed the identity of the Texel

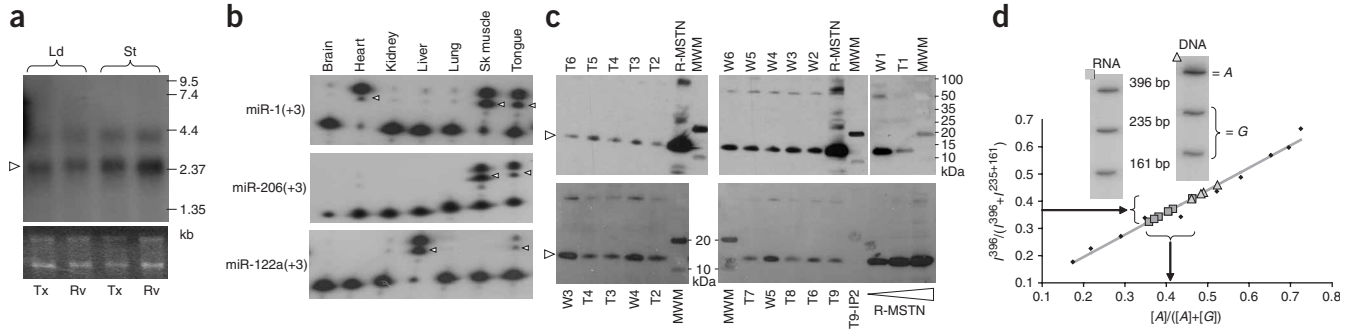

Figure 2 Expression analysis of GDF8 and potentially interacting miRNAs. (**a**) RNA blot analysis of *GDF8* expression in longissimus dorsi (Ld) and semitendinosus (St) of Texels (Tx) and wild-type Romanovs (Rv). Arrow: *GDF8* transcript. The ethidium bromide–stained gel is shown. (**b**) Expression analysis of *miR-1*, *miR-122a* and *miR-206* in adult sheep tissues. Arrowheads: products of the primer extension. Numbers in parentheses: expected number of nucleotides added by reverse transcriptase, given the sequence of primers and mature miRNA. (**c**) Detection of circulating myostatin by immunoprecipitation (IP) and protein blotting in Texels (T1–T9) and controls (W1–W6) (matched for sex (male) and age (4 months)). Arrowheads: 12.5-kDa band corresponding to mature myostatin monomer. Except for T1 and W1, all samples were analyzed blindly. T9-IP2: second IP performed on sample T9 showing that the IP depletes myostatin from the sample. R-MSTN: recombinant myostatin[11]. The three R-MSTN lanes in the lower panel correspond to recombinant myostatin at 25, 62.5 and 312.5 ng/ml. MWM: molecular weight markers. (**d**) Estimating the relative amounts of *A* (Texel) versus *G* (wild-type) transcripts in *AG* heterozygotes. The *y* axis shows relative intensity of the 396-bp fragment (*A* allele) compared with the sum of the intensities of the 396-bp and 235 + 161 bp (*G* allele) fragments; the *x* axis predicts the proportion of *A* molecules. Diamonds: data points obtained by mixing varying amounts of genomic DNA from *AA* and *GG* animals, yielding the gray calibration curve. Triangles: results obtained with *AG* genomic DNA. Squares: results obtained with skeletal muscle RNA from *AG* heterozygotes.

and control mRNA sequences. At first glance, these results suggested that Texel animals produce normal levels of functional *GDF8* mRNA.

However, in light of our fine-mapping results, we decided to further examine *GDF8*. We sequenced 10.5 kb spanning the sheep *GDF8* gene from the same three Texel animals and seven controls. This identified 20 SNPs (**Supplementary Fig. 4** online). None of these reside in particularly conserved sequence elements. We genotyped all the SNPs on 42 Texels, 90 controls (11 breeds) and the four *TR* rams (three F1 animals and F2 20254) (**Supplementary Table 2** online).

The first notable observation was the virtual monomorphism of Texels, contrasting with the considerable variation of the 11 other breeds. This would be expected if a *GDF8* mutation had undergone a selective sweep in Texels.

We were able to exclude 18 SNPs because at least one of the four *TR* rams was homozygous for the SNP. For these SNPs, the Texel allele was at high frequency in the other breeds ($\geq$0.62; **Supplementary Table 2**). Two SNPs could not be excluded on the same basis: *g-2449C-G*, located 2.5 kb upstream from the transcription start site, and *g+6723G-A*, located in the 3′ UTR. For *g-2449C-G*, the frequency of the *C* allele was 98% in Texels versus 11.5% in controls. For *g+6723G-A*, the frequency of the *A* allele was 99% in Texels versus 1% in controls; thus, the *g+6723G-A A* allele seemed to be virtually Texel-specific.

g-2449C-G is located in a region devoid of conserved sequence elements. It is difficult to imagine how it might affect myostatin function, especially given the comparable amounts of intact *GDF8* mRNA found in Texels and controls. Moreover, one of the Texel animals was heterozygous only for *g-2449C-G* (**Supplementary Table 2**) and homozygous for the Texel allele for all other SNPs. The easiest explanation is that this animal is *TT* but that it inherited one *T* haplotype with a recombination just upstream of the *GDF8* gene. This would exclude *g-2449C-G*.

Closer examination of the sequences flanking *g+6723G-A* showed that the *A* allele creates one of the 3′ UTR octamer motifs (ACATTCCA) discovered earlier[8] and assumed to correspond to

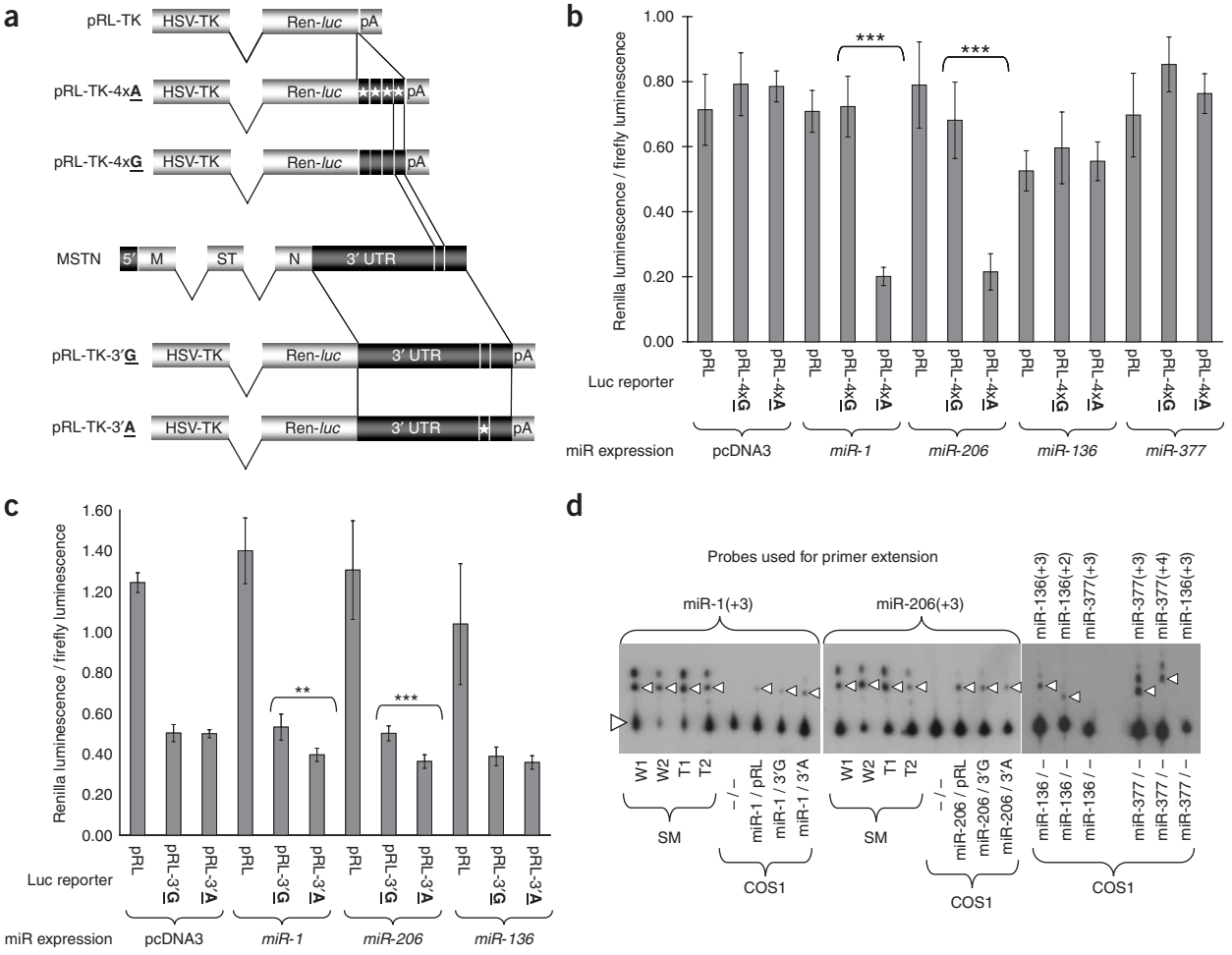

Figure 3 Reporter assay testing the interaction between miRNA-*GDF8* interaction. (**a**) Schematic representation of the Renilla luciferase expression vectors and the *GDF8* gene (MSTN). (**b**,**c**) Renilla-to-firefly luminescence ratios observed when cotransfecting COS1 cells with the indicated luciferase reporter and either an empty pcDNA3 vector or a pcDNA3 vector expressing the miRNAs as indicated (*miR-1, miR-206, miR-136* and *miR-377*). Error bars: ± 1.96 s.e.m. obtained from three (**b**) or five (**c**) replicates. ***: $P < 0.001$; **: $P < 0.01$. (**d**) Detection of the mature miRNAs in sheep skeletal muscle (SM) and in transfected COS1 cells. Probes are given above the autoradiograms and the origin of the RNA samples underneath. W1–W2 correspond to SM RNA from two controls; T1–T2 to SM RNA from two Texels. In the case of the COS1 RNA, we mention the corresponding pcDNA3 and pRL-TK vectors used for transfection. Small arrowheads mark the positions of the expected primer extension products; the large arrowhead marks the position of the unextended probes. Equal amounts (5 μg) of total RNA were used in all cases.

miRNA targets. The probability that a random mutation in the sheep *GDF8* 3′ UTR creates one of these 540 octamer motifs is 0.045; the probability to create an octamer motif with an equally high conservation score (18.2; ref. 8) is 0.0088. This suggested that *g+6723G-A* might contribute to the muscular hypertrophy by causing miRNA-mediated translational inhibition of the Texel *GDF8* allele.

Three known miRNAs target the ACATTCCA octamer described above: *miR-1*, *miR-206* and *miR-122a* (ref. 8). Notably, *miR-1* has previously been shown to be strongly expressed in skeletal muscle and heart in the mouse[9]. We designed primer pairs based on interspecies alignments to amplify the sheep miRNA orthologs. Note that in man and mice, mature *miR-1* is processed from two paralogous genes: *miR-1.1* and *miR-1.2*. We obtained PCR products for the four miRNA genes and sequenced them. This confirmed the presence of the corresponding genes in the sheep and the perfect conservation of the mature miRNAs, including their seed (**Supplementary Fig. 5** online). We examined their expression profiles in sheep by primer extension (**Fig. 2b**). Although *miR-122* was expressed in liver and tongue but not in skeletal muscle, *miR-1* and *miR-206* were strongly expressed in skeletal muscle and tongue. As in the mouse, we also detected *miR-1* in the heart. The observation of a strong expression of *miR-1* and *miR-206* in skeletal muscle, the primary site of *GDF8* expression, lent further strength to our hypothesis.

If *g+6723G-A* causes miRNA-mediated translational repression of *GDF8* transcripts, Texel sheep should have reduced levels of circulating myostatin. Indeed, myostatin protein is detected in serum of wild-type mice and humans but is absent in serum of mice and humans homozygous for *GDF8* loss-of-function mutations[10,11]. Using a monoclonal antibody to myostatin, we immunoprecipitated myostatin from serum of nine Texel animals and six controls and detected its presence using a polyclonal antibody specific for the mature protein by protein blotting. We identified a 12.5-kDa band corresponding to mature myostatin in all 15 sera (**Fig. 2c**). Most importantly, the intensity of the 12.5-kDa band in Texel animals was approximately one third of the intensity in wild-type animals.

The model of miRNA-mediated translational inhibition predicts reduced stability for the mutant *GDF8* transcript owing to accelerated degradation in P-bodies (see, for example, ref. 12). To test this, we compared the relative abundance of *A* versus *G* transcripts in skeletal muscle of *AG* heterozygotes. This approach is more sensitive, as it is less affected by variation between individuals and between samples, as is observed when comparing *GDF8* levels between homozygotes. The wild-type *G* transcripts were indeed ~1.5 times more abundant than the *A* transcripts (**Fig. 2d**).

We then aimed to test the interaction between mutant *GDF8* transcripts and *miR-1* and *miR-206* directly. We cloned four tandem repeats[9] of an 80-bp sequence centered around *g+6723G-A* into the 3′ UTR of luciferase reporter vectors (creating constructs pRL-TK-4×A, containing Texel sequences, and pRL-TK-4×G, containing control sequences (**Fig. 3a**)). We cotransfected COS1 cells with these reporter constructs and with pcDNA3 vectors expressing either *miR-1* and *miR-206* or the control *miR-136* and *miR-377*. In agreement with our prediction, when we cotransfected the reporter constructs with pcDNA3 vectors expressing either *miR-1* or *miR-206*, we observed a highly significant ($P = 0.0002$ and 0.0011, respectively) reduction of pRL-TK-4×A signal to ~30% of the signal obtained with pRL-TK-4×G or the unmodified pRL-TK (**Fig. 3b**). On the other hand, when we cotransfected the same reporter constructs with an empty pcDNA3 vector or with pcDNA3 vectors expressing *miR-136* or *miR-377*, there was no significant difference between luminescence obtained with pRL-TK-4×A, pRL-TK-4×G or unmodified pRL-TK.

We repeated these experiments using luciferase vectors in the 3′ UTR of which we cloned the complete 1.5-kb mutant or wild-type 3′ UTR, creating constructs pRL-TK-3′A and pRL-TK-3′G, respectively (**Fig. 3a**). Inserting the *GDF8* 3′ UTR in the vector reduced the luminescence to ~40% of the signal obtained with the unmodified vector. More notably, the signal obtained from pRL-TK-3′A and pRL-TK-3′G did not differ when we cotransfected the COS1 cells with an empty pcDNA vector or with one expressing *miR-136*, whereas we observed a significant reduction of the pRL-TK-3′A signal to ~70% of the pRL-TK-3′G signal when expressing either *miR-1* or *miR-206* ($P = 0.0029$ and 0.0003, respectively) (**Fig. 3c**).

In these experiments, we were able to detect mature miRNAs only after transfection with the cognate pcDNA3 vectors. The expression levels of *miR-1* and *miR-206* in COS1 cells were, however, lower than in sheep muscle (**Fig. 3d**).

Our results support a model in which a point mutation in the *GDF8* 3′ UTR creates an illegitimate target site for at least two miRNAs that are strongly expressed in skeletal muscle. This results in miRNA-mediated translational downregulation and reduction in myostatin concentrations contributing to muscular hypertrophy. It is tempting to speculate that such hypomorphic alleles are not a sheep *GDF8* idiosyncrasy but that they make a substantial contribution to phenotypic variation in other organisms, including man. To evaluate the frequency of polymorphic miRNA-target interactions, we searched among 73,497 SNPs mapping to the 3′ UTR of 13,621 human genes for those creating or destroying one of the octamers from ref. 8. The ancestral allele was determined from the orthologous chimpanzee sequence. We identified 2,490 putative Texel-like SNPs, creating an illegitimate miRNA target site. In addition, we identified 2,597 SNPs destroying at least one motif. Of these 2,597 SNPs, 483 affect an octamer perfectly conserved across the four analyzed mammalian species. These 483 have a fairly high likelihood of affecting gene regulation and hence phenotypic variation. Among SNPs uncovering target sites, those promoting miRNA-anti-target interactions[13] are of most interest. They could be identified by comparing the expression profile of target and miRNA. That polymorphic miRNA-target interactions may contribute to disease is illustrated by the recent identification of a SNP associated with Tourette syndrome that affects the interaction between *miR-189* and the 3′ UTR of *SLITRK1* (ref. 14). We have performed the same analyses for the mouse, using the orthologous rat sequences to infer ancestral state for 77,283 SNPs in the 3′ UTR of 10,200 genes. We have identified 1,182 SNPs creating and 1,321 SNPs destroying putative miRNA target sites, 234 of which are evolutionary conserved. We have generated a website (Patrocles) that lists SNPs that have the potential to affect miRNA-target interactions.

The nature of the mutations that underlie genetic variation for complex traits remains a matter of debate: do quantitative trait nucleotides (QTN) primarily affect gene structure or gene regulation? Are epistatic interactions between QTN the rule or the exception? This work identifies a new class of regulatory mutations that might make an important contribution to the heritability of complex traits. It also points toward possible epistatic interactions between polymorphisms in miRNA genes and their targets.

METHODS

Map construction. Primers for the amplification of microsatellite markers were obtained from public-domain cattle and sheep maps. Microsatellite genotyping was performed using standard procedures. Linkage maps were constructed using CRIMAP[15].

QTL mapping. QTL mapping was performed using QTL Express[16]. The nominal *P* values of the *F* statistics generated by QTL Express were Bonferroni

corrected for 17 independent tests (as deduced from the permutations performed by QTL Express) to obtain chromosome-wide *P* values and then were Bonferroni corrected for 27 additional tests (corresponding to the number of sheep chromosomes) to obtain genome-wide *P* values. Confidence intervals for the QTL location were obtained by bootstrapping implemented in QTL Express.

Marker-assisted segregation analysis. The likelihoods of the pedigree data were computed as

$$L = \prod_i^{n_{\mathrm{rec}}} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(Ph_i-(M+\frac{a}{2}))^2}{2\sigma^2}} \prod_j^{n_{\mathrm{T}}} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(Ph_j-(M-\frac{a}{2}))^2}{2\sigma^2}}$$

where n_{rec} represents the number of offspring inheriting the recombinant Texel-Romanov chromosome from the sire (see **Fig. 1b**), and n_{T} represents the number of offspring inheriting the nonrecombinant Texel chromosome from their sire. Ph_i and Ph_j correspond to the phenotypic values of the i^{th} and j^{th} offspring, respectively; *M* represents the midpoint between the means of the *rec* and *T* offspring; *a* refers to the $R \rightarrow T$ QTL allele substitution effect and σ^2 corresponds to the residual variance. To compute the likelihood of the data assuming that the sire is homozygous *TT* for the QTL (L_{TT}), we set *a* at 0 and estimated the values of *M* and σ^2 that maximized the likelihood. To compute the likelihood of the data assuming that the sire is heterozygous *TR* for the QTL (L_{TR}), we jointly estimated the values of *a*, *M* and σ^2 that yielded the highest likelihood of the data (L_{ML}). $2\mathrm{LN}(L_{ML}/L_{TT})$ was assumed to have a χ^2 distribution with one degree of freedom under the null *TT* hypothesis.

Selective sweep detection. To detect the effects of a putative selective sweep on the allelic frequency spectrum in hypermuscled Texel animals compared with control animals, we analyzed the microsatellite genotypes of 41 hypermuscled Texel animals and 108 wild-type controls representing 16 different breeds (Blackbelly: 2; Booroola: 4; Ile de France: 12; Lacaune: 5; Merinos: 2; Rambouillet: 10; Romanov: 17; Southdown: 3; Suffolk: 10; Tarasconnaise: 9; Targee: 3; Berrichon du Cher: 6; Blanc du Massif Central: 5; Charmoise: 4; Charollais: 8; Manech: 8) using DISLAMB[5].

Resequencing of the myostatin gene. To resequence the *GDF8* gene from Texel and control animals we (i) amplified the coding parts of the *GDF8* gene from genomic DNA in three PCR products of 372, 375 and 381 bp, respectively; (ii) amplified the entire *GDF8* ORF by RT-PCR from skeletal muscle mRNA in two overlapping PCR products of 805 and 625 bp, respectively and (iii) amplified 10.5 kb spanning the *GDF8* gene from genomic DNA in 13 overlapping segments. The corresponding primer pairs are listed in **Supplementary Table 3** online. For RT-PCR, RNA was extracted from skeletal muscle using Trizol (Invitrogen), and cDNA was synthesized using the SuperScript III First Strand Synthesis System for RT-PCR (Invitrogen). All PCR products were gel purified using Geneclean (Qbiogene) and sequenced on both strands using the same primers and BigDye Terminator v3.1 Cycle Sequencing kits (Applied Biosystems) and a 3730 DNA Analyzer (Applied Biosystems).

RNA blot analysis. Total RNA was extracted from skeletal muscle using RNA Insta-Pure (Eurogentec), and mRNA was isolated using the Oligotex Direct mRNA mini kit (Qiagen). The mRNA was size fractionated using the Reliant RNA Gel System (BMA) and was blotted on an Ambion membrane in 5× SSC, 10 mM NaOH. The membrane was hybridized in Ultrahyb (Ambion) buffer at 42 °C to a sheep *GDF8* cDNA probe labeled with the Random-Primed DNA Labeling Kit (Boehringer Mannheim). The membrane was washed at 42 °C in 2× SSC, 0.1% SDS (wt/vol), washed in 0.1× SSC, 0.1% SDS and subjected to autoradiography.

Genotyping of the g+*6723G-A* and other myostatin SNPs. Genotyping of the g+*6723G-A* SNP was done by PCR–restriction fragment length polymorphism analysis. A 1,003-bp fragment encompassing the SNP was amplified by PCR from genomic DNA using primers g+6723G-A.UP1 and g+6723G-A.DN1 (**Supplementary Table 3**), digested using *Hpy*CH4IV and size fractionated by agarose gel electrophoresis. The *g+6723G-A* SNP destroys a restriction site that cleaves the *G* allele (but not the *A* allele) into a 270-bp and a 733-bp fragment. To genotype the 19 other *GDF8* SNPs, we amplified six amplicons from genomic DNA using standard PCR conditions and primers reported in **Supplementary Table 3**, purified them using the Multiscreen PCR µ96 Filter Plate (Millipore) and sequenced them using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and a 3730 DNA Analyzer (Applied Biosystems).

PCR amplification and sequencing of the sheep *miR-1.1*, *miR-1.2*, *miR-122* and *miR-206* genes. The human, mouse, rat and cattle *miR-1.1*, *miR-1.2*, *miR-122* and *miR-206* genes were aligned using ClustalW. Primer pairs (**Supplementary Table 3**) were designed in conserved segments of the gene and used to amplify the orthologous sheep genes by PCR from genomic DNA. The PCR products were gel purified using Geneclean (Qbiogene) and sequenced on both strands using the same primers and a BigDye Terminator v3.1 Cycle Sequencing Kit and a 3730 DNA Analyzer.

Primer extension assay to detect mature miRNAs. miRNA expression was evaluated by primer extension as previously described[17] using primers reported in **Supplementary Table 3**.

Detection of myostatin protein by immunoprecipitation and protein blotting. We prepared beads coupled to a monoclonal antibody (JA-16) to myostatin (directed against a C-terminal peptide of myostatin), and we immunoprecipitated myostatin by incubating 60 µl packed beads with ~0.4 ml of serum[18]. Serum volumes were adjusted for total protein concentration, which was determined using the Quick Start Bradford Protein Assay (Biorad). After washing, bound myostatin was eluted with Laemmli buffer. Samples were separated by SDS-PAGE, blotted on a nitrocellulose membrane and probed with rabbit polyclonal antibody to myostatin (L8014).

Measuring *g+6723G-A* allelic imbalance at the mRNA level using hot-stop PCR. Total RNA was extracted from skeletal muscle of heterozygous *AG* animals (three 70-d-old fetuses and four 4-month-old animals) using Trizol (Invitrogen). The RNA was treated with TurboDNase (Ambion). cDNA was synthesized using SuperScriptIII First Strand Synthesis System for RT-PCR (Invitrogen). Hot-stop PCR was performed according to ref. 19, using the g+6723G-A.UP2 and g+6723G-A.DN2 primers (**Supplementary Table 3**), amplifying a 396-bp fragment of the *GDF8* 3′ UTR. The labeled PCR products were digested with *Hpy*CH4IV (cleaving the *G* allele into a 235-bp and a 161-bp fragment) and were size fractionated by denaturing PAGE. The intensity of the respective restriction fragments were quantified using a Phosphorimager (Molecular Dynamics). The proportion of *A* allele was estimated from the ratio $I^{396} / (I^{396} + I^{235+161})$ (where I^x corresponds to the intensity of the corresponding fragment) and from a calibration curve established using template DNA with known *A*-to-*G* ratios (see **Fig. 2d**).

Testing the interaction between the Texel myostatin 3′ UTR and *miR-1* and *miR-206* using a dual-luciferase reporter assay in COS1 cells. To construct the pRL-TK-4×A (Texel) and pRL-TK-4×G (wild-type) vectors, we amplified an 80-bp fragment of the *GDF8* 3′ UTR encompassing the *g+6723G-A* SNP from genomic DNA of a Texel animal and a Romanov animal using primers Xba-ovmyo1211-f (with an *Xba*I tail) and Spe-ovmyo1290-r (with an *Spe*I tail) (**Supplementary Table 3**). The primers were chosen to avoid the occurrence in the final construct of secondary RNA structures that might occlude the miRNA target site as assessed using RNAfold[20]. *Xba*I- and *Spe*I-digested (New England Biolabs) PCR products were self-ligated (LigaFast Rapid DNA Ligation System, Promega). *Xba*I/*Spe*I-resistant tetramers were gel purified and ligated in the *Xba*I site of the pRL-TK vector (Promega). To construct the pRL-TK-3′A (Texel) and pRL-TK-3′G (wild-type) vectors, we amplified the entire *GDF8* 3′ UTR from genomic DNA of a Texel animal and a Romanov animal using primers Xba-ovmyo3′UTR-f (with an *Xba*I tail) and Spe-ovmyo3′UTR-r (with an *Spe*I tail) (**Supplementary Table 3**). The *Xba*I- and *Spe*I-digested PCR products were cloned in the *Xba*I site of the pRL-TK vector. Plasmid DNA was purified using the EndoFree plasmid maxi kit (Qiagen), and the inserts of all constructs were completely sequenced. The sheep *miR-1.1*, *miR-206*, *miR-136* and *miR377* genes were amplified from genomic DNA using primer pairs with *Hin*dIII and *Nhe*I tails (**Supplementary Table 3**). The *Hin*dIII/*Nhe*I-digested

PCR products were directionally cloned in the *Hin*dIII/*Nhe*I site of the pcDNA3.1(+) vector (Invitrogen). COS1 cells (European Collection of Cell Cultures (ECACC) no. 88031701) were maintained in DMEM supplemented with 10% fetal bovine serum, 2 mM glutamine, 0.1 mM non-essential amino acids, penicillin (100 units/ml) and streptomycin (100 µg/ml). Using Lipofectamine 2000 (Invitrogen) following the manufacturer's recommendations, we transfected the 0.8×10^5 cells per well in 24-well plates with a mixture comprising 400 ng of pRL-TK Renilla luciferase construct, 400 ng of pcDNA3.1 construct and 10 ng of pGL3 firefly luciferase control vector (Promega). The luciferase assays were performed 24 h after transfection using the dual-luciferase reporter assay system (Promega) and a Centro LB960 luminometer (Berthold Technologies).

URLs. dbSNP: http://www.ncbi.nlm.nih.gov/SNP/. Patrocles database, listing SNPs that have the potential to affect miRNA-target interactions: http://www.patrocles.org.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This project was supported by grants from the (i) Walloon Ministry of Agriculture (D31/1036), (ii) the 'GAME' Action de Recherche Concertée from the Communauté Française de Belgique, (iii) the Interuniversity Attraction Pole P5/25 from the Belgian Federal Science Policy Office (R.SSTC.0135), (iv) the European Union 'Callimir' Specific Targeted Research Project (STREP), (v) the University of Liège, (vi) the French Research Agency Genanimal, (vii) Région Auvergne + Départements INRA de Génétique Animale + CEPIA (Caractérisation et élaboration des produits issus de l'agriculture) and (viii) the Région Limousin and Université de Limoges. A.C. benefited from a fellowship of the Département de Génétique Animale, INRA. A.C. and H.T. both benefit from an E.U. Marie-Curie postdoctoral fellowship. C.C. is a 'chercheur qualifié' from the Fonds National de la Recherche Scientifique. We are grateful to P. Leroy and H. Leveziel for their continuous interest and support for this work; to the technical personnel at the Langlade experimental station; to Labogena and France-Upra-Sélection for providing us with DNA samples; to the Centre d'Insémination de Faulx-les-Tombes in Belgium; to C. Fasquelle, V. Marot and V. Dhennin for technical assistance; to J. Vandessompele for advice with real-time PCR and to M. Van Poucke for the primer sequences of the endogenous controls.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Marcq, F. *et al.* Preliminary results of a whole-genome scan targeting QTL for carcass traits in a Texel-Romanov intercross. *Proc. 7th World Cong. On Genetics Appl. Livest. Prod., Montpellier, France* 323–326 (2002).
2. Laville, E. *et al.* Effects of a QTL for muscle hypertrophy from Belgian texel sheep on carcass conformation and muscularity. *J. Anim. Sci.* **82**, 3128–3137 (2004).
3. Haley, C.S., Knott, S.A. & Elsen, J.-M. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195–1207 (1994).
4. Johnson, P.L., McEwan, J.C., Dodds, K.G., Purchas, R.W. & Blair, H.T. A directed search in the region of GDF8 for quantitative trait loci affecting carcass traits in Texel sheep. *J. Anim. Sci.* **83**, 1988–2000 (2005).
5. Terwilliger, J.D. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic loci. *Am. J. Hum. Genet.* **56**, 777–787 (1995).
6. Lee, S.J. & McPherron, A.C. Myostatin and the control of skeletal muscle mass. *Curr. Opin. Genet. Dev.* **9**, 604–607 (1999).
7. Tobin, J.F. & Celeste, A.J. Myostatin, a negative regulator of muscle mass: implications for muscle degenerative diseases. *Curr. Opin. Pharmacol.* **5**, 328–332 (2005).
8. Xie, X. *et al.* Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature* **434**, 338–345 (2005).
9. Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214–220 (2005).
10. Zimmers, T.A. *et al.* Induction of cachexia in mice by systematically administered myostatin. *Science* **296**, 1486–1488 (2002).
11. Schuelke, M. *et al.* Myostatin mutation associated with gross muscle hypertrophy in a child. *N. Engl. J. Med.* **350**, 2682–2688 (2004).
12. Zamore, P.D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309**, 1519–1524 (2005).
13. Bartel, D.P. & Chen, C.-Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* **5**, 396–400 (2004).
14. Abelson, J.F. *et al.* Sequence variants in *SLITRK1* are associated with Tourette's syndrome. *Science* **310**, 317–320 (2005).
15. Lander, E. & Green, P. Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA* **84**, 2363–2367 (1987).
16. Seaton, G., Haley, C.S., Knott, S.A., Kearsey, M. & Visscher, P.M. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339–340 (2002).
17. Davis, E. *et al.* RNAi-mediated allelic trans-interaction at the imprinted callipyge locus. *Curr. Biol.* **15**, 743–749 (2005).
18. Hill, J.J. *et al.* The myostatin propeptide and the follistatin-related gene are inhibitory binding proteins of myostatin in normal serum. *J. Biol. Chem.* **277**, 40735–40741 (2002).
19. Uejima, H., Lee, M.P., Cui, H. & Feinberg, A.P. Hot-stop PCR: a simple and general assay for linear quantitation of allele ratios. *Nat. Genet.* **25**, 375–376 (2000).
20. Hofacker, I.L. *et al.* Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.* **125**, 167–188 (1994).