© 1997 Nature Publishing Group http://www.nature.com/naturegenetics

A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle

Luc Grobet[1], Luis José Royo Martin[2], Dominique Poncelet[1], Dimitri Pirottin[1], Benoit Brouwers[1], Juliette Riquet[1], Andreina Schoeberlein[1], Susana Dunner[2], François Ménissier[3], Julio Massabanda[4], Ruedi Fries[4], Roger Hanset[1] & Michel Georges[1]

An exceptional muscle development commonly referred to as 'double-muscled' (Fig. 1) has been seen in several cattle breeds and has attracted considerable attention from beef producers. Double-muscled animals are characterized by an increase in muscle mass of about 20%, due to general skeletal-muscle hyperplasia—that is, an increase in the number of muscle fibers rather than in their individual diameter[1]. Although the hereditary nature of the double-muscled condition was recognized early on, the precise mode of inheritance has remained controversial; monogenic (dominant and recessive), oligogenic and polygenic models have been proposed[2]. In the Belgian Blue cattle breed (BBCB)[4], segregation analysis performed both in experimental crosses[3] and in the outbred population suggested an autosomal recessive inheritance. This was confirmed when the muscular hypertrophy (*mh*) locus was mapped 3.1 cM from microsatellite TGLA44 on the centromeric end of bovine chromosome 2 (ref. 5). We used a positional candidate approach to demonstrate that a mutation in bovine *MSTN*, which encodes myostatin, a member of the TGFβ superfamily, is responsible for the double-muscled phenotype. We report an 11-bp deletion in the coding sequence for the bioactive carboxy-terminal domain of the protein causing the muscular hypertrophy observed in Belgian Blue cattle.

Previous work has shown that the pro-a(III) collagen gene (*Col3AI*) is located in the same chromosomal region as the *mh* locus. *Col3AI* has been mapped to BTA2q12–22 by *in situ* hybridization[6], while a *Col3AI* RFLP marker was shown to be closely linked to TGLA44 (θ=2%)[7]. This identifies the region flanking *Col3AI* on the human map—HSA2q31–33—as the likely orthologous human chromosome segment, which is compatible with Zoo-FISH data[6] and mapping of type-I markers on somatic cell hybrids[8].

To refine the correspondence between the HSA2q31–33 and BTA2q12–22 maps, we developed comparative anchored tagged sequences (CATS)—primer pairs that would amplify an STS from the orthologous gene in different species[9]—for a series of genes flanking *Col3A1* on the human map and for which sequence information was available in more than one mammal. In addition to *Col3AI*, CATS were obtained for *Col5A2*, *INPP1*, *TFPI*, *TTN*, *CHN*, *GAD1*, *CTLA4* and *CD28*. We used these CATS, as well as all microsatellite markers available for proximal BTA2—TGLA44, BM81124, BM3627, ILSTS026, INRA40 and TGLA431 (ref. 10)—to screen a six-genome equivalent bovine YAC library (A. Schoeberlein, manuscript in preparation). We explored potential overlap between YACs obtained with this panel of STSs on the basis of common STS content, as well as cross-hybridization between individual SINE-PCR product. This analysis revealed three independent YAC contigs in the region of interest (Fig. 2): i) contig A, containing microsatellites TGLA44, BM81124 and *INPP1*; ii) contig B, containing *Col3AI* and *Col5A2;* and iii) contig C, containing microsatellite markers BM3627, ILSTS026, INRA40 and *TFPI*. We isolated new polymorphic microsatellite markers from contig B (BULGE20), as well as contig A (BULGE23), in order to increase the local information content. We genotyped a previously described experimental backcross population[5,15], and used all available genotypes to construct a linkage map with the ILINK program. We obtained the following most likely order and sex-averaged recombination rates between adja-

Fig. 1 Double-muscled Belgian Blue animal homozygous for the *nt821del(11)* deletion in the myostatin gene.

Table 1 • Distribution of *nt821(del11)* mutation in different cattle breeds

Breed	Phenotype	Genotype		
		+/+	+/nt821(del11)	nt821(del11)/nt821(del11)
Belgian Blue	DM	–	–	104
Asturiana	DM	–	–	52
Maine-Anjou	DM	42	–	–
Holstein-Friesian	Normal	100	–	–
Jersey	Normal	96	–	–

[1]Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 20 Bd de Colonster, 4000 Liège, Belgium. [2]Laboratorio de Genética, Dpto. de Produción Animal, Universidad Complutense, 28040 Madrid, Spain. [3]Station de Génétique Quantitative et Appliquée, I.N.R.A., 78352 Jouy-en-Josas, France. [4]Technische Universität München, D-85350 Freising-Weihenstephan, Germany. Correspondence should be addressed to M.G. e-mail: michel@stat.fmv.ulg.ac.be

© 1997 Nature Publishing Group http://www.nature.com/naturegenetics

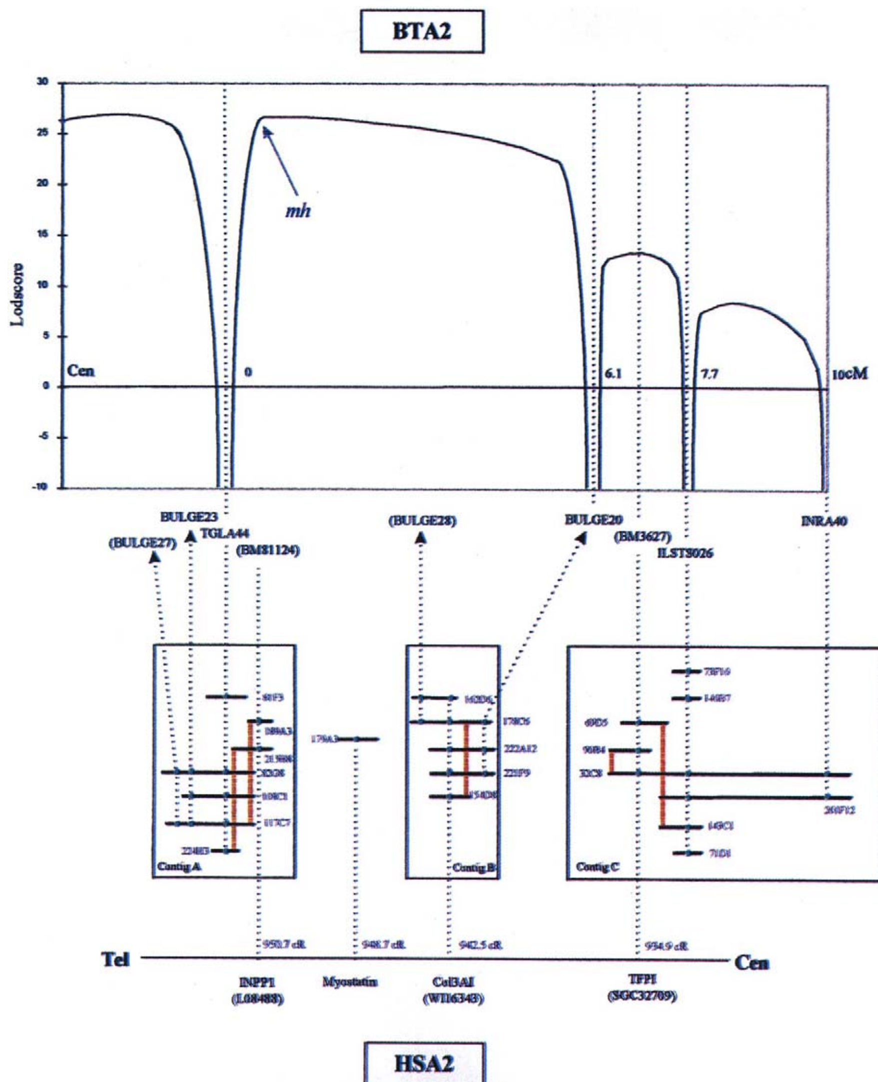

Fig. 2 Summary of genetic, physical and comparative mapping data around the bovine *mh* locus. A multipoint lod-score curve obtained for the *mh* locus with respect to the microsatellite marker map is shown. Markers that were not informative in our pedigree are shown between brackets; their map position is inferred from published mapping data. Markers and the YACs from which they were isolated are connected by arrows. The RH map of the relevant section of human HSA2 is shown, with the relative position in cR of the ESTs used in this work. Stippled lines connect microsatellite and type I markers with their respective positive YACs. YACs showing cross-hybridizing SINE-PCR products are connected by the red boxes.

cent markers: [TGLA44-(0%)-BULGE23]-(6,1%)-BULGE20-(1,6%)-ILSTS026-(2.3%)-INRA40-7,1%)-TGLA431. The position of BULGE20 between TGLA44 and ILSTS026 confirmed the order of the three contigs, which had been predicted from available human mapping data (Fig. 2).

The *mh* locus was placed on this new marker map by multipoint linkage analysis, yielding a lod-score curve (Fig. 2) and positioning the *mh* locus in the TGLA44-BULGE20 interval, with an associated maximum lod score of 26.4. Given the relative position of these microsatellite markers with respect to *INPP1* and *Col3AI*, as deduced from the integration of the human and bovine map, these results indicated that the *mh* gene was likely to be located in a chromosome segment bounded by *INPP1* and *Col3AI*.

Recently, McPherron *et al.*[11] demonstrated that mice homozygous for a knockout deletion of *Mstn* (formerly called *Gdf8*), encoding myostatin, were characterized by a general increase in skeletal-muscle mass, similar to the bovine double-muscled phenotype. Using the published murine *Mstn* cDNA sequence, we identified a tentative human consensus (THC) cluster in the Unigene database representing three cDNA clones (221299, 300367, 308202) and six EST sequences (H92027, H92028, N80248, N95327, W07375, W24782). The corresponding THC covered 1,296 bp of human cDNA, showing a homology of 78.1% with murine myostatin when averaged over the entire sequence and 91.1% when only the translated parts of the human and murine genes (566 bp) were considered. This THC therefore clearly corresponded to the human orthologue of the murine myostatin gene. We designed primers to amplify a 272-bp fragment from the second exon of the human myostatin gene (*MSTN)* and to genotype the Genebridge-4 radiation hybrid panel[12]. We placed *MSTN* on the Whitehead/MIT framework radiation hybrid map, at position 948.7 cR of the HSA2 map with an associated lod score greater than 3. Closer examination of the *MSTN* segregation vector and comparison with the vectors from all markers located in that region showed it to be identical to EST SGC38239, placed on the Whitehead/MIT radiation hybrid map[13] at position 946.8 cR of HSA2. This places *MSTN* in the interval between *Col3AI* (EST WI16343–942.5 cR) and *INPP1* (EST L08488–950.2–951.2 cR; Fig. 2), making bovine *MSTV* a very strong positional candidate for the *mh* gene.

To test whether myostatin is associated with the double-muscled phenotype in cattle, we designed primer pairs based on the available mouse and human sequences, aiming for RT-PCR amplification of the entire coding sequence from bovine cDNA. We identified two primer pairs that jointly amplified the entire coding sequence as well as 76 bp and 83 bp of 5′ and 3′ untranslated sequence, respectively, in two overlapping DNA fragments of 753 (primers GDF8.27–GDF8.12) and 724 bp (primers GDF8.11–GDF8.21) long. We successfully amplified the expected DNA products from cDNA generated from skeletal muscle of both a wild-type (homozygous +/+) and a double-muscled (homozygous *mh/mh*) animal, and sequenced them on both strands.

The nucleotide sequence of the wild-type allele shows 89.1% identity with the mouse *Mstn* coding sequence over a 1,128-bp overlap. The predicted protein shows 92.5% identity in a 375-amino-acid overlap with mouse myostatin. As expected for a member of the *TGFβ* superfamily, the bovine myostatin gene is characterized by a proteolytic processing site thought to mediate cleavage of the bioactive carboxy-terminal domain from the longer N-terminal fragment, and by nine cysteine residues separated by a characteristic spacing and believed to be involved in intra- and inter-molecular disulphide bridges[14].

The nucleotide sequence obtained from the *mh* allele was identical to the wild-type allele, except for an 11-bp deletion in nucleotides

© 1997 Nature Publishing Group http://www.nature.com/naturegenetics

Table 2 • Primer pairs for PCR amplification

CATS		
INPP1	5′-CAGCAAAGTCCTTAATGGTAACAAGC-3′	5′-GGGTCACTGAAGAAAACGTCCTG-3′
COL3A1	5′-CCCCATATTATGGAGATGAACCG-3′	5′-AGTTCAGGATGGCAGAATTTCAG-3′
COL5A2	5′-GCAAACTGGGYGGRAGCAAGACC-3′	5′-TTSTTCCTGGGCTTTTATTGAGAC-3′
TFPI	5′-AAGCCWGATTTCTGCTTYTTGGAAG-3′	5′-TGCCMAGGCAHCCRCCRTACTTGAA-3′
TTN	5′-GGTCGTCCTACACCAGAAG-3′	5′-GGTTGACATTGTCAAGAACAAG-3′
CHN	5′-TCTCMAAAGTCGTCTGTGACAATC-3′	5′-TGYTCRTTTTCTTTCAGAGTTGC-3′
GAD1	5′-RCTGGTCCTCTTCACCTCAGAAC-3′	5′-ACATTGTCVGTTCCAAAGCCAAG-3′
CTLA4	5′-AGGTYCGGGTGACDGTGCTKC-3′	5′-TGGRTACATGAGYTCCACCTTGC-3′
CD28	5′- AGCTGCARGTATWCCTACAAYCT-3′	5′-GTYCCRTTGCTCYCTCTCRTTGYC-3′
Microsatellite markers		
TGLA44	5′-AACTGTATATTGAGAGCCTACCATG-3′	5′-CACACCTTAGCGACTAAACCACCA-3′
BULGE27	5′- CTACCTAACAGAATGATTTTGTAAG-3′	5′-AGTGTTCTTGCCTAGAGAATCCCAG-3′
BULGE23	5′-ACATTCTCTCACCAATATGACATAC- 3′	5′-TAAGTCACCATTACATCCTTAGAAC-3′
BM81124	5′-GCTGTAAGAATCTTCATTAAGCACT- 3′	5′-CCTGATACATGCTAAGGTTAAAAAC-3′
BULGE28	5′-AGGCATACATCTGGAGAGAAACATG-3′	5′-CAGAGGAGCCTAGCAGGCTACCGTC-3′
BULGE20	5′-CAGCAGGTCTGTTGAAGTGTATCAG-3′	5′-AGTGGTAGCATTCACAGGTAGCCAG-3′
BM3627	5′-CAGTCCATGGCACCATAAAG-3′	5′-TCCGTTAGTACTGGCTAATTGC-3′
ILSTS026	5′-CTGAATTGGCTCCAAAGGCC-3′	5′-AAACAGAAGTCCAGGGCTGC-3′
INRA40	5′- TCAGTCTCCAGGAGAGAAAAC-3′	5′-CTCTGCCCTGGGGATGATTG-3′
Bovine myostatin primers		
MSTN	5′- TTTCACTTGGCATTACTCAAAAG-3′	5′-CAATGCCTAAGTTGGATTCAGGTTG-3′
MSTN	5′-ACAGTGTTTGTGCAAATCCTGAGAC-3′	5′-ATACTCWAGGCCTAYAGCCTGTGGT-3′

821–831 (counting from the initiation codon). This frameshifting deletion, occurring after the first cysteine residue of the carboxy-terminal domain, drastically disrupts the downstream amino-acid sequence and reveals a premature stop-codon at codon 14 after the deletion (Fig. 3). This mutation virtually eliminates the bioactive part of the molecule and is therefore likely to be the cause of the recessive double-muscled phenotype. Following conventional nomenclature, this mutation will be referred to as *nt821(del11)*.

To support our theory that this mutation causes the double-muscled condition, we designed primer pairs flanking the deletion (Fig. 3), and genotyped our experimental backcross population[5]. Correspondence between phenotype and genotype was perfect for the entire pedigree. All ten BBCB double-muscled sires were homozygous for the *nt821(del11)* mutation and all 41 F1 females were heterozygous, while the deletion segregated in the backcross generation as predicted from the phenotypes of the backcross animals: all 53 double-muscled offspring were homozygous for the mutation, and the remaining 55 normal animals were heterozygous.

To examine the distribution of the *nt821(del11)* mutation in different populations, we genotyped a cohort of 196 normal animals representing two dairy breeds (Holstein-Friesian and Jersey) and a cohort of 198 double-muscled animals representing three beef breeds: BBCB, Asturiana and Maine-Anjou (Table 1). The role of the *mh* locus in

Fig. 3 ***a***, Electropherograms obtained by cycle sequencing of the myostatin cDNA sequence from a double-muscled and a normal animal, showing the *nt821del(11)* deletion in the double-muscled animal. The primers used to amplify the fragment encompassing the deletion from genomic DNA are shown in green. ***b***, Amino-acid sequence of the murine (green), bovine normal (black) and bovine *nt821del(11)* (blue) allele. The bovine normal allele is chosen as master sequence; identity with the master sequence in the murine and bovine *nt821del(11)* allele is indicated by dashes. The putative site of proteolytic processing is boxed in red, while the nine conserved cysteines in the carboxy-terminal region are boxed in black.

© 1997 Nature Publishing Group http://www.nature.com/naturegenetics

BBCB[5] and Asturiana[15] was reported previously. Linkage analysis, performed in Maine-Anjou pedigrees segregating for the double-muscled phenotype, yielded a maximum lod score of 5.2 at 6.4% recombination from TGLA44, clearly involving the same chromosomal region and probably the *mh* gene in this breed as well. All 196 dairy animals were homozygous for the wild-type allele. In BBCB and Asturiana, all double-muscled animals were homozygous for the *nt821(del11)* deletion, demonstrating allelic homogeneity in these two breeds. In contrast, all double-muscled Maine-Anjou were homozygous for the wild-type allele and did not carry the *nt821(del11)* deletion. The most likely explanation for this finding would be allelic heterogeneity of the bovine myostatin gene. If confirmed, this would force re-evaluation of the popular hypothesis that a unique mutation was disseminated across Europe with the once popular Shorthorn breed at the beginning of the 19th century[2]. Allelic heterogeneity in the bovine myostatin gene might also account for the conflicting genetic models proposed for the inheritance of the trait in different breeds. As members of the TGFβ superfamily are known to act as dimers, dominant negative mutations could account for the dominant inheritance observed for the double-muscled condition in some breeds.

The identification of *MSTN* as the gene causing the double-muscled phenotype will allow for the development of diagnostic tests that will facilitate the selection for or against this trait in cattle. It also paves the way towards manipulation of muscle development in livestock by means of various approaches, including immunomodulation and transgenesis. It would be interesting to determine whether inactivation of *MSTN* after birth, using a *Cre-loxP* system or related approach, might still lead to increased muscle development. Moreover, the identification of the myostatin gene as a key regulator of muscle development will permit study of upstream and downstream factors (such as the myostatin receptor) that might lead to the identification of other genes underlying genetic variation for muscle development in livestock.

Methods

Genetic mapping. Genetic mapping was performed in a previously described (Holstein-Friesian x Belgian Blue) x Belgian Blue experimental backcross population with 108 informative offspring[5]. The pedigree material available in Maine-Anjou was composed of five paternal half-sib pedigrees comprising a total of 42 double-muscled offspring from five conventional sires. Microsatellite genotyping was performed according to standard procedures[16], using the primer sequences reported in Table 2. Linkage analyses were performed with the MLINK, ILINK and LINKMAP programs of LINKAGE (version 5.1; ref. 17) and FASTLINK (2.3P version, June 1995; ref. 18) packages. Linkage analysis of the *mh* locus was performed under a simple recessive model assuming full penetrance for *mh/mh* individuals and zero penetrance for the two other genotypes.

Screening of the YAC library. The yeast artificial chromosome library was screened by PCR a three-dimensional pooling scheme as previously described[19]. The primer pairs corresponding to the CATS used to screen the library are reported in Table 2. Cross-hybridization between SINE-PCR products of individual YACs was performed as described (using primers reported in ref. 21)[20]. Microsatellites were isolated from YACs as described[22].

Mapping of the human myostatin gene. DNA from the Genebridge-4 panel[12] was purchased from Research Genetics (Huntsville, AL) and genotyped by PCR according to standard procedures and the following primer pair corresponding to the human myostatin gene sequence: 5′–GGCCCAACTATGGATATATTTG–3′ and 5′–GGTCCTGGGAAGGTTACAGCA–3′. Mapping was performed with the WWW server of the Whitehead Institute/MIT Center for Genome Research and their *RH-mapper* program (Slonim, D., Stein, L., Kruglyak, L. & Lander, E., unpublished data) to position the markers with respect to the framework map. Segregation vectors of the query markers were compared with the vectors from all markers in the region of interest in the complete Data Release 11.9 (May 1997) to obtain a more precise position.

RT-PCR. Total RNA was extracted from skeletal muscle (triceps brachialis) as described[23]. RT-PCR was performed with the Gene-Amp RNA PCR Kit (Perkin Elmer) and the primers reported in Table 2. The PCR products were purified with QiaQuick PCR Purification kit (Qiagen) and sequenced with Dye terminator Cycle Sequencing Ready Reaction (Perkin Elmer) and an ABI373 automatic sequencer, using the primers reported in Table 2.

Acknowledgements

This work was supported by a grant from the Belgian Ministère de l'Agriculture et des Classes Moyennes (D1/2-5744A) and a grant from the European Union (B104-CT95-0073). We are grateful to J. Weissenbach and C. Fizames for communicating the segregation vector of the L08488 EST in the GeneBridge-4 panel, and to D. Serteyn, I. Dufrasne, D. Desmecht and F. Rolin for helping us to obtain the muscle biopsies. We acknowledge Holland Genetics and Livestock Improvement Corporation for providing us with the samples from Holstein-Friesian and Jersey animals. We thank P. Leroy for his continuous support and for funding the construction of the YAC library.

Received 2 July 1997; accepted 14 July 1997.

1. Hanset, R. The major gene of muscular hypertrophy in the Belgian Blue cattle breed. in *Breeding for Disease Resistance in Farm Animals* (eds Owen, J.B. & Axford, R.F.E.) 467–478 (C.A.B. International, Oxford, 1991).
2. Ménissier, F. Present state of knowledge about the genetic determination of muscular hypertrophy or the double muscled trait in cattle. in *Current Topics in Veterinary Medicine and Animal Science*, vol 16: *Muscle Hyertrophy of Genetic Origin and Its Use to Improve Beef Production* (eds King, J.W.B. & Ménissier, F.) 387–428 (Martinus Nijhoff, Norwell, Massachusetts, 1982).
3. Hanset, R. & Michaux, C. On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. I. Experimental data. *Genet. Sel. Evol.* **17**, 359–368 (1985).
4. Hanset, R. & Michaux, C. On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed: II. Population data. *Genet. Sel. Evol.* **17**, 369–386 (1985).
5. Charlier, C. *et al.* The *mh* gene causing double-muscling in cattle maps to bovine chromosome 2. *Mamm. Genome* **6**, 788–792 (1995).
6. Solinas-Toldo, S., Lengauer, C. & Fries, R. Comparative genome map of man and cattle. *Genomics* **27**, 489–496 (1995).
7. Fisher, S.R., Beever, J.E. & Lewin, H.A. Genetic mapping of COL3AI to bovine chromosome 2. *Mamm. Genome* **8**, 76–77 (1996).
8. O'Brien, S.J. *et al.* Anchored reference loci for comparative genome mapping in mammals. *Nature Genet.* **3**, 103–112 (1993).
9. Lyons, A.L. *et al.* Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nature Genet.* **15**, 47–56 (1996).
10. Kappes, S.M. *et al.* A second-generation linkage map of the bovine genome. *Genome Res.* **7**, 235–249 (1997).
11. McPherron, A.C., Lawler, A.M. & Lee, S.-J. Regulation of skeletal muscle mass in mice by a new TGFβ superfamily member. *Nature* **387**, 83–90 (1997).
12. Walter, M.A., Spillett, D.J., Thomas, P., Weissenbach, J. & Goodfellow, P.N. A method for constructing radiation hybrid maps of whole genomes. *Nature Genet.* **7**, 22–28 (1994).
13. Hudson T.G. *et al.* An STS-based map of the human genome. *Science* **270**, 1945–1954 (1995), with supplementary data from the Whitehead Institute/MIT Center for Genome Research, Human Genetic Mapping Project, data release 11.9 (May 1997).
14. McPherron, A.C. & Lee, S.-J. The transforming growth factor β superfamily. in *Growth Factors and Cytokines in Health and Disease*, vol 1B (eds LeRoith, D. & Bondy, C.) 357–393 (Jai Publishing, Greenwich, Connecticut, 1996).
15. Dunner, S. *et al.* Towards interbreed IBD fine mapping of the *mh* locus: double-muscling in the Asturiana de los Valles breed involves the same locus as in the Belgian Blue cattle breed. *Mamm. Genome* **8**, 430–435 (1997).
16. Georges, M. *et al.* Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* **139**, 907–920 (1995).
17. Lathrop, M. & Lalouel, J.M. Easy calculations of lodscores and genetic risk on small computers. *Am. J. Hum. Genet.* **36**, 460–465 (1984).
18. Cottingham, R.W., Jr., Idury, R.M. & Schäffer, A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* **53**, 252–263 (1993).
19. Libert, F., Lefort, A., Okimoto, R., Womack, J. & Georges, M. Construction of a bovine genomic library of large yeast artificial chromosome clones. *Genomics* **18**, 270–276 (1993).
20. Hunter, K.W. *et al.* Toward the construction of integrated physical and genetic maps of the mouse genome using interspersed repetitive sequence PCR (IRS–PCR) genomics. *Genome Res.* **6**, 290–299 (1996).
21. Lenstra, J.A., van Boxtel, J.A.F., Zwaagstra, K.A. & Schwerin, M. Short interspersed nuclear element (SINE) sequences of the *Bovidae*. *Anim. Genet.* **24**, 33–39 (1993).
22. Cornélis, F. *et al.* Identification of a CA repeat at the TCRA locus using yeast artificial chromosomes: a general method for generating highly polymorphic markers at chosen loci. *Genomics* **13**, 820–825 (1992).
23. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299 (1979).