Haplotype Sharing Refines the Location of an Imprinted Quantitative Trait Locus With Major Effect on Muscle Mass to a 250-kb Chromosome Segment Containing the Porcine *IGF2* Gene

Carine Nezer,* Catherine Collette,* Laurence Moreau,* Benoît Brouwers,* Jong-Joo Kim,* Elisabetta Giuffra,† Nadine Buys,‡ Leif Andersson§ and Michel Georges*,[1]

**Department of Genetics, Faculty of Veterinary Medicine, University of Liège, 4000-Liège, Belgium, §Department of Medical Biochemistry and Microbiology, Uppsala University, S-751 24 Uppsala, Sweden, †Centro Ricerche Studi Agroalimentari FPTP-CERSA, LITA, 20090 Segrate, Italy and ‡Gentec, 9255 Buggenhout, Belgium*

Manuscript received November 28, 2002
Accepted for publication May 13, 2003

ABSTRACT

We herein describe the fine mapping of an imprinted QTL with major effect on muscle mass that was previously assigned to distal SSC2p in the pig. The proposed approach exploits linkage disequilibrium in combination with QTL genotyping by marker-assisted segregation analysis. By identifying a haplotype shared by all "*Q*" chromosomes, we map the QTL to an ~250-kb chromosome segment containing *INS* and *IGF2* as the only known paternally expressed genes. This considerably reinforces the candidacy of these genes, justifying their detailed analysis.

QUANTITATIVE trait loci (QTL) mapping has become a preferred approach toward the molecular dissection of quantitative traits, whether of fundamental, medical, or agronomic importance. A multitude of chromosomal locations predicted to harbor genes influencing traits of interest have been identified using this strategy (*e.g.*, Andersson 2001; Flint and Mott 2001; Mackay 2001; Mauricio 2001). In most cases, however, the mapping resolution is in the order of the tens of centimorgans, which is insufficient for positional cloning of the underlying genes. High-resolution mapping of QTL therefore remains one of the major challenges in the genetic analysis of complex traits.

Three factors limit the achievable mapping resolution: marker density, crossover density, and the ability to deduce QTL genotype from phenotype. Increasing marker density may still be time-consuming in many organisms but is conceptually the simplest bottleneck to resolve. Two options are available to increase the local crossover density: breed recombinants *de novo* or exploit historical recombination events; *i.e.*, use linkage disequilibrium (LD). The former approach is generally used with model organisms that have a short generation interval (*e.g.*, Darvasi 1998), while the latter is the only practical alternative when working with human or large livestock species. Optimal use of LD to fine map QTL in outbred populations is presently an area of very active research (*e.g.*, Ardlie *et al.* 2002). The ability to deduce QTL genotype from phenotype can be improved by using "clones" (*e.g.*, recombinant inbred lines; Darvasi 1998), by means of progeny testing (*e.g.*, Georges *et al.* 1995), or by marker-assisted segregation analysis (*e.g.*, Riquet *et al.* 1999).

Recently a QTL with major effect on muscle mass and fat deposition was mapped to the distal end of chromosome arm SSC2p in the pig (Jeon *et al.* 1999; Nezer *et al.* 1999). The most likely position of the QTL was shown to coincide with a chromosome region that is orthologous to HSA11p15 in the human, which is known to harbor an imprinted domain. QTL analyses performed with imprinting models strongly suggested that the underlying gene was indeed imprinted and expressed only from the paternal allele (Jeon *et al.* 1999; Nezer *et al.* 1999). The human 11p15 imprinted domain is known to contain at least 9 imprinted transcripts. Three of these are paternally expressed: *LIT-1* (*KVLQT1-AS*), *IGF2*, and *IGF2-AS* (*e.g.*, Reik and Walter 2001). Fifteen imprinted transcripts are known to map to the orthologous domain on distal mouse chromosome MMU7, of which 4 are paternally expressed: *Lit-1* (*Kvlqt1-as*), *Ins2*, *Igf2*, and *Igf2-as* (*e.g.*, http://www.mgu.har.mrc.ac.uk/imprinting/imprinting.html; Onyango *et al.* 2000). Because of its known function in myogenesis (Florini *et al.* 1996), *IGF2* stood out as a prime positional candidate. However, no sequence variations that could account for the observed QTL effect were found in the coding parts of the porcine *IGF2* gene (Jeon *et al.* 1999; Nezer *et al.* 1999).

To refine the map position of this QTL and to verify whether its position remained compatible with a direct role of the *IGF2* gene, we applied an approach targeting the three factors limiting the mapping resolution of QTL: (i) we increased the marker density of the chromo-

[1]*Corresponding author:* Department of Genetics, Faculty of Veterinary Medicine, University of Liège (B43), 20 Bd de Colonster, 4000-Liège, Belgium. E-mail: michel.georges@ulg.ac.be

some region of interest; (ii) we determined the QTL genotype of a number of individuals by marker-assisted segregation analysis; and (iii) we applied a LD-based haplotype-sharing approach to determine the most likely position of the QTL. This approach is analogous to the one that was previously applied by Riquet *et al.* (1999) to refine the map position of a QTL influencing milk production in dairy cattle. It makes the assumption that the observed QTL effect is due to the segregation of a QTL allele with major substitution effect ("*Q*") that appeared by mutation or migration *g* generations ago and swept through the populations as a result of artificial selection. As a consequence, at the present generation, *n* chromosomes carrying the *Q* allele are expected to share a haplotype of size $\sim 2/ng$ (in morgans) containing the QTL (Dunner *et al.* 1997).

By doing so we have identified a shared haplotype spanning ~250 kb that is predicted to contain the quantitative trait nucleotide (QTN; Mackay 2001). The corresponding chromosome segment contains *INS* and *IGF2* as the only known paternally expressed genes. This considerably enforces the candidacy of these two genes and demonstrates that LD can be exploited to map QTL in outbred populations to chromosome intervals containing no more than a handful of genes.

MATERIALS AND METHODS

Pedigree material and phenotypic data: The pedigree material used for this work was composed of a subset of a previously described Piétrain × Large White F_2 pedigree (Hanset *et al.* 1995; Nezer *et al.* 2002), as well as a series of paternal half-sib pedigrees sampled in commercial lines derived from the Piétrain and Large White breeds (N. Buys, personal communication). In the F_2 animals, percentage of lean meat was measured as "percentage of lean cuts" as previously described (Hanset *et al.* 1995), while in the commercial lines percentage of lean meat was measured as "Piglog" corresponding to (63.6882 − 0.4465 a − 0.5096 b + 0.1281 c), where a is millimeters of backfat measured between the third and fourth lumbar vertebra at 7 cm from the spine, b is millimeters of backfat measured between the third and fourth last rib at 7 cm from the spine, and c is millimeters of loin thickness, measured at same position as b.

Marker-assisted segregation analysis: The QTL genotype of each sire was determined from a *Z*-score, corresponding to the $\log_{10}$ of the likelihood ratio L_{H_1}/L_{H_0}, where L_{H_1} corresponds to the likelihood of the pedigree data assuming that the boar is of *Qq* genotype, and L_{H_0} corresponds to the likelihood of the pedigree data assuming that the boar is of *QQ* or *qq* genotype. The corresponding likelihoods were computed as

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{(-(y_i-(\bar{y}+0.5a))^2)/2\sigma^2} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\sigma}} e^{(-(y_j-(\bar{y}-0.5a))^2)/2\sigma^2}.$$

In this *n* is the number of informative offspring having inherited the "left" homolog from their sire, *m* is the number of informative offspring having inherited the "right" homolog from their sire, $y_{i(j)}$ is the phenotype of offspring i (j), $\bar{y}$ is the average phenotype of the corresponding pedigree computed over all (informative and noninformative) offspring, σ is the residual standard deviation maximizing *L*, and *a* is the *Q* to *q* allele substitution effect. $|a|$ was set at zero when computing L_{H_0} and at 2.0% when computing L_{H_1} (Nezer *et al.* 1999). The sign of *a* (+ or −) was selected to maximize L_{H_1}. Boars were considered to be *Qq* when $Z > 2$, *QQ* or *qq* when $Z < -2$, and of undetermined genotype if $2 > Z > -2$.

Linkage disequilibrium analysis: Probabilities for two chromosomes to be identical by descent (IBD) at a given map position conditional on flanking marker data were computed according to Meuwissen and Goddard (2001). The effective population size (N_e) was set at 200 on the basis of estimates of N_e determined from LD data (N. Harmegnies, unpublished observations) and the number of generations to the base population at 20. A multipoint test for association was performed using the DISMULT program described in Terwilliger (1995).

RESULTS

QTL genotyping by marker-assisted segregation analysis: We genotyped a series of paternal half-sib families, counting at least 20 offspring for two microsatellite markers located on the distal end of chromosome arm SSC2p: *SWR2516* and *SWC9* (Jeon *et al.* 1999; Nezer *et al.* 1999). These families originated either from a previously described Piétrain × Large White F_2 pedigree (Nezer *et al.* 2002) or from two composite pig lines derived from Large White and Piétrain founder animals (N. Buys, personal communication).

The pedigrees from sires that were heterozygous for one or both of these markers were kept for further analysis. Twenty such pedigrees could be identified for a total of 941 animals. Offspring were sorted in three classes on the basis of their marker genotype: "L" (left homolog inherited from the sire), "R" (right homolog inherited from the sire), or "?" (not informative or recombinant in the *SWR2516-SWC9* interval).

Offspring were slaughtered at a constant weight of ~105 kg, and a series of phenotypes were collected on the carcasses, including percentage of lean meat, measured as either percentage of lean cuts (experimental cross) or Piglog (composite lines; see materials and methods).

We then computed the likelihood of each sire family under two hypotheses: H_0, postulating that the corresponding boar was homozygous at the QTL, and H_1 postulating that the boar was heterozygous at the QTL. Assuming a biallelic QTL, H_0 corresponds to QTL genotypes *QQ* or *qq*, and H_1 to genotype *Qq*. Likelihoods were computed using percentage of lean meat as phenotype (as the effect of the QTL was shown to be most pronounced on this trait in previous analyses) and assuming a *Q* to *q* allele substitution effect of 2.0% (Nezer *et al.* 1999). If the odds in favor of one of the hypotheses were superior or equal to 100:1, the most likely hypothesis was considered to be true (see materials and methods).

Using these rules, we could determine the QTL genotype for 14 of the 20 boars. Seven of these proved to be heterozygous *Qq* and the other 7 to be homozygous and thus of either *QQ* or *qq* genotype (Figure 1).

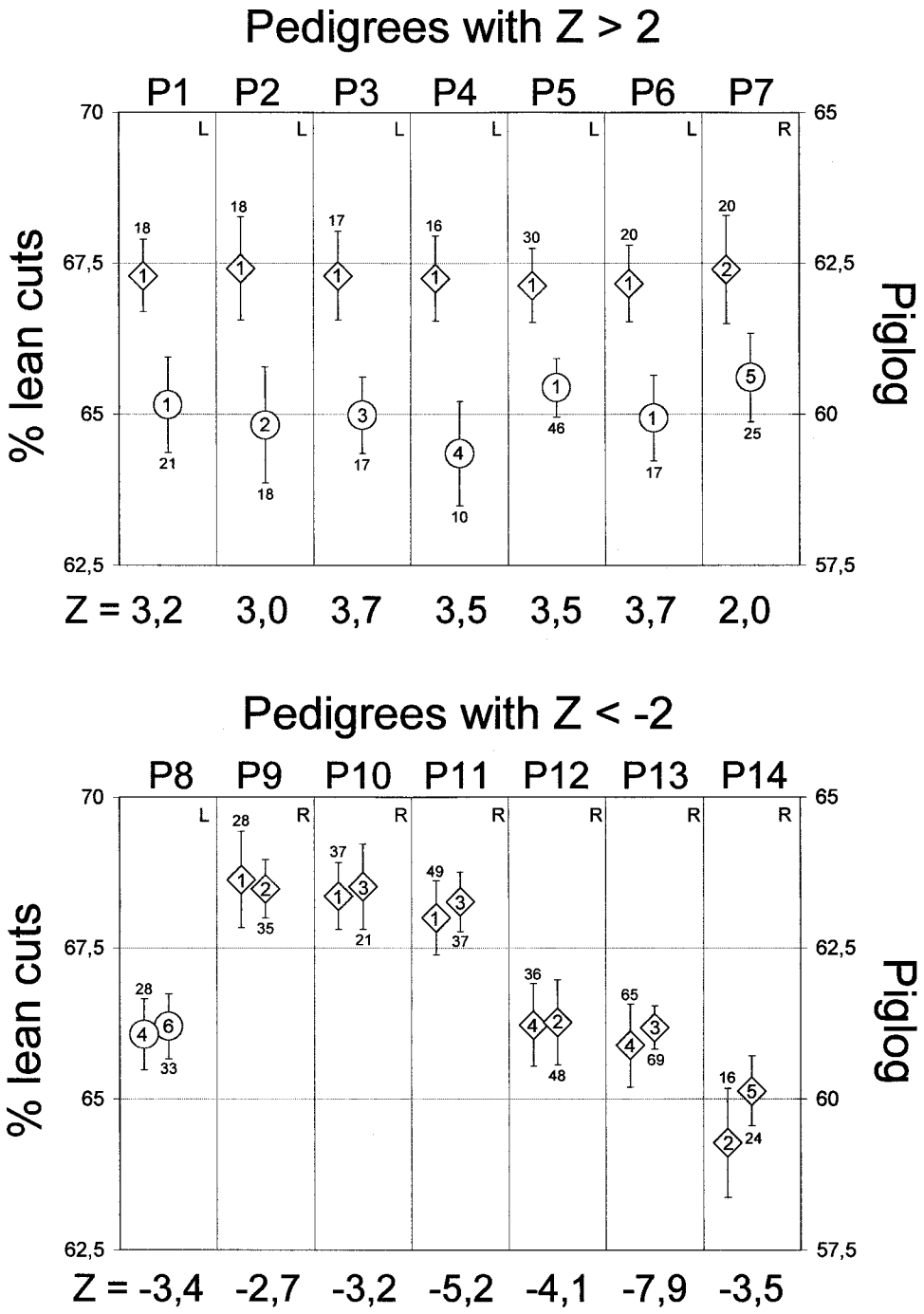

FIGURE 1.—QTL genotyping by marker-assisted segregation analysis. The graphs show, for 14 paternal half-sib pedigrees (P1, P2, . . . , P14), the phenotypic mean ±2 standard errors of the offspring sorted in two groups according to the homolog inherited from the sire. The number of offspring in each group is given above and under the error bars, respectively. The top graph corresponds to the boars that were shown to be heterozygous *Qq* for the QTL and the bottom graph to the boars that were shown to be homozygous at the QTL. Pedigrees for which the percentage of lean meat was measured as percentage of lean cuts (NEZER *et al.* 2002) are marked by L (left axis), those for which "Piglog" was used (see MATERIALS AND METHODS) are marked by R (right axis). The graph reports a *Z*-score for each pedigree, *i.e.*, the $\log_{10}$ of the H_1/H_0 likelihood ratio where H_1 assumes that the boar is heterozygous *Qq* for the QTL, while H_0 assumes that the boar is homozygous *QQ* or *qq*. *Q* alleles associated with a positive allele substitution effect on percentage of lean meat are marked by a diamond, *q* alleles by a circle. The number within the symbols differentiates the *Q* and *q* alleles according to the associated marker haplotype (see RESULTS and Figure 2).

Constructing a physical and genetic map of the porcine ortholog of the human 11p15 imprinted domain: We developed porcine sequence-tagged sites (STS) across the orthologous region of the human 11p15 imprinted domain. The *SWC9* marker was known from previous studies to correspond to a $(CA)_n$ microsatellite located in the 3′ untranslated region (UTR) of the porcine *IGF2* gene (JEON *et al.* 1999; NEZER *et al.* 1999; AMARGER *et al.* 2002). Fourteen novel STS were developed in genes [*TSSC5, KVLQT1* (3×), *CD81, TH* (2×), *INS* (3×), *IGF2* (3×), and *H19*] and 5 in intergenic regions [*IG(IGF2-H19), IG(H19-RL23mrp)* (4×)]. The corresponding primer sequences were derived from the porcine genomic sequence when available (AMARGER *et al.* 2002) or from porcine expressed sequence tags (EST) that were identified by BLAST searches using the human orthologs as query sequences (Table 1).

We screened a porcine bacterial artificial chromosome (BAC) library (FAHRENKRUG *et al.* 2001) by filter hybridization, using (i) human cDNA clones corresponding to genes know to map to 11p15, as well as (ii) some of the 19 previously described porcine STS. Six of the identified BACs were shown by PCR to contain at least one of the porcine STS available in the region

TABLE 1

Utilized sequence-tagged sites (STS) and corresponding DNA sequence polymorphisms (DSP)

STS	Source	UP-primer (5′-3′)	DN-primer (5′-3′)	DSP[b]	DSP (5′-3′)
TSSC5(I1)[a]	BI183986	TCATCCAGGGCCTGGTCATCG	TGTCTGAGGCCGACACGGCC	T1	CCCCCTCCC(C/T)GGCCCCC
				T2	ACCCAGGGC(C/T)CCTTGAG
KVLQT1(SSR)	BAC956B11	CTTTGAGGTCCATCATGTTCCA	GGACGTACATCCCATCGATGA	SSR	$(CA)_n$
KVLQT1(I12)	BF198846	ATGGTTGTCCTCTGCGTGGGC	TGGCGGTCGACGTGCAGCATC	T1	TGGGTGGGGG(C/T)GCAGCCCC
				T2	GCTGGGA(C/T)CAGACC(G/A)TCTGGG
				T3	GCTGGGA(C/T)CAGACC(G/A)TCTGGG
				T4	CTGTCTGCTCAT(C/T)CGGGGGCTG
				T5	GGCTGCGGGAGC(C/T)TGGGGCCAC
				T6	GCCACCCCCGCC(C/T)TGACCCTGA
KVLQT1(I11)	BF198846	ATCCGCTTCCTCCAGATCCTG	GCCGATGTACAGCGTGGTGA	V1	TCTGGGCCGG(G/T)GTCCCCG
				T1	AAAAGGGTCC(A/G)GGAAGCT
				T2	TTGCAAACAGC(C/T)CCCAGAAGG
				T3	AGAAGGCGCAG(C/T)CTCCACGGG
				T4	AGGGCGCTGG(C/T)TGCAGGGGTG
				T5	TTTATGAGTC(A/G)CAAAAACGAG
				T6	TGATGTCCGCC(C/T)(G/T)GGCAGACT
				V2	TGATGTCCGCC(C/T)(G/T)GGCAGACT
KVLQT1(I7)	BF198846	GCCCCAAGCCCAAGAAGTCTG	CCAGAATTGTCACAGCCATCC	T	TCCGGGGCAT(A/G)TAGGACTGG
389B2T7	BAC389B2	GGAGTACCTGCTGTGGCTTAGTG	CGTCCTATATCCATCAGGAATATTG	T1	AGTAGTAT(C/T)CATGAGCAC
				T2	CCCAGGCCTC(G/A)ATCAGCTGGTTG
				V	TATATGCCA(C/A)ACATGTGGCCCT
CD81(I3)	F23061	GGGGCCATCCAGGAGTCACAG	CAAAGAGGATCACGAGGCAGG		
370SNP15	370C17	CTGAGATGTTAGACTGGGTTG	CCCAGCATCTCCTCACGGTC	ID	GCCGGTGGTC(C)AGGGCTTGTGC
370SNP6	370C17	TCCCTCTCAGCTGCGGTTTGG	AATCCAGGGAAAACTGCTCGTC	T	CTCTCACTGG(C/T)TGAGCC
370SNP1	370C17	ATGGGTGCGCTGTTAATTTGAAC	CCCCCTGGGCCTTCTGGG	T1	TACACAC(A/G)CGTGAAA
				T2	ACTCGTGC(A/G)TTTTAG
				V	TTTTAGGTAA(A/C)TTACACC
370SNP2	370C17	TCCTCCCCGCCCCCCTCC	CATTGTCACAGTTGAAATGCGG	T1	CCCCAG(G/A)GGGCTACA
				T2	TTTGCTTGC(A/G)GTGCGGCT
				T3	CCCGGCCG(C/T)CTGCAGCCT
INRA370SP6	BACINRA370	TGCGTAGCCATGGCGATGGGG	AGTGTGGAACCCTGGGGGGGGAAGG		
370C17T7	BAC370C17	AGAGGGTACAGAAGCCCTG	TTTGGTGTGGTGTCTGCTGACCC		
PULGE3	BACINRA370	AGGCTTTCTATC TGCAGGAGG	ACCGTGTGGCCATCTGGGTG	SSR	TCTCTGTAT$(CA)_n$CGCACGCAC
PULGE1	BACINRA370	GCGTTGCAGTGGCTCTGGCG	GACACGGCCGCATGAATGTGC	SSR	ACCCCAACA$(TA)_n$ATTATGGTA
TH(I13A)	AY044828	GCCCGTCTACTTCGTGTCTGAG	ATCTCTGCCTTCATCGCACCCCC	V	AGGATCCAGCC(A/T) GCAGCCCCG
				ID	TCACAACCCCC(C) TCCCACAGC
				T	CTGCGGAGGGG(A/G)GACCTGCAG
TH(I13B)	AY044828	GCTGCGGACCCCACCGTCAC	AGACTTCACCCCTAAAAGCCTGG	ID	GCCAGGT(CAAGGCCAGGT)CGAGGCC
INS(5′)	AY044828	AGCAGGCTGCTGTGCTGGG	AGCCCAGACCCAGCTGACGG	T1	GGCGCTTATGG(G/A)GCCGGGAGC
				V	CAAGCCCGG(G/T)CGGTTTGGCCT
				T2	CTAATGACCTC(A/G)AGGCCCCCA

(*continued*)

TABLE 1

(Continued)

STS	Source	UP-primer (5′-3′)	DN-primer (5′-3′)	DSP[b]	DSP (5′-3′)
INS(I1,E2,I2)	AY044828	TGATGACCCACGGAGATGATCC	GCAGTAGTTCTCCAGCTGGTAGAGGGAA	T1	GGGACCAGCTG(C/T)GTTCCCAGG
				V	GCCCTGCTGGC(C/G)CTCTGGGCG
				T2	CTCCCACGCCC(C/T)GGTCCCGCT
INS(3′)	AY044828	GCTCTCGCCACATCGGCTGC	GGCGCCCAGCTCTAGGCCCGGC	T	GGGCTGGCTGC(G/A)GTCTGGGAG
IGF2(E3)	AY044828	CCCCTGAACTTGAGGACGAGCAGCC	CGCTGTGGGCTGGGTGGGCTGCC	T	GCTGCCCCCCA(A/G)CCTGAGCTG
IGF2(E5)	AY044828	CTTGCCTCCAACTCCCTCCC	AGTGAACGTGAAACGGGGGG	SSR	CTCTC GCT GTC $(CT)_n$ CGCCCTCTCTT
IGF2(I8)	AY044828	TGCGCCACCCCCGCCAAGTCC	GCTTCCAGGTGTCATAGCGGAAG	V	AGCCGGCTCCT(G/C)GGCTTCAAG
				T	AGAGGTTGTTG(C/T)TCTGGGACA
SWC9	AY044828	AAGCACCTGTACCCACACG	GGCTCAGGGATCCCACAG	SSR	$(CA)_n$
IG(IGF2-H19)	AY044828	CAAGCCAGGTCCTGTCGAGG	GGACCCTGGGGGCTGTGG	T	CGGCCTGTGGC(A/G)GGGAAGCTG
H19	AY044828	ACGGTCCCGGGTCAGCAGG	CAGAGCAAGTGGGCACCCAG	T1	CGCGGGTTTGG(C/T)CAGCGGCAG
				T2	CACAGAGGACA(C/T)GGCCGCTTC
				T3	TCCTGGGGGCC(C/T)GCGGCTCGT
IG(H19-RL23mrp)A	AY044828	GAGCACAGCCAAAGAACGGCCG	CTTCACCCACGGACATGGCCGC	T	CACCCAGGCTG(C/T)GCCCTGCGT
IG(H19-RL23mrp)B	AY044828	CGGGGGCACTGGGGGTCC	CCGAGACCCTCCTCAAGTCC	T	GTTCGCCCTCC(A/G)CTCTCAGCA
IG(H19-RL23mrp)C	AY044828	TGAGCTGCTGAGCCCACAGG	CAAGGGAAAGGTGTGCCGACC	T	GGCCGGGCGCT(C/T)CGCCTTCCC
IG(H19-RL23mrp)D	AY044828	AGGCAGAGGGCAGAGAGGGG	CTCCAGCCCCACACTCTGC	T	GCGTCCAGCGC(C/T)GAATCAGGC
SWR2516	gi7643973	GTGCATTATCGGGAGGTATG	ACCCTGTATGATACTGTAACTCTGG	SSR	ATAGGGTTA$(GT)_n$AGATCAGTC

[a] I, intron; E, exon.
[b] DSP, type of DNA sequence polymorphism: T, transition; V, transversion; ID, insertion/deletion; SSR, simple sequence repeat.

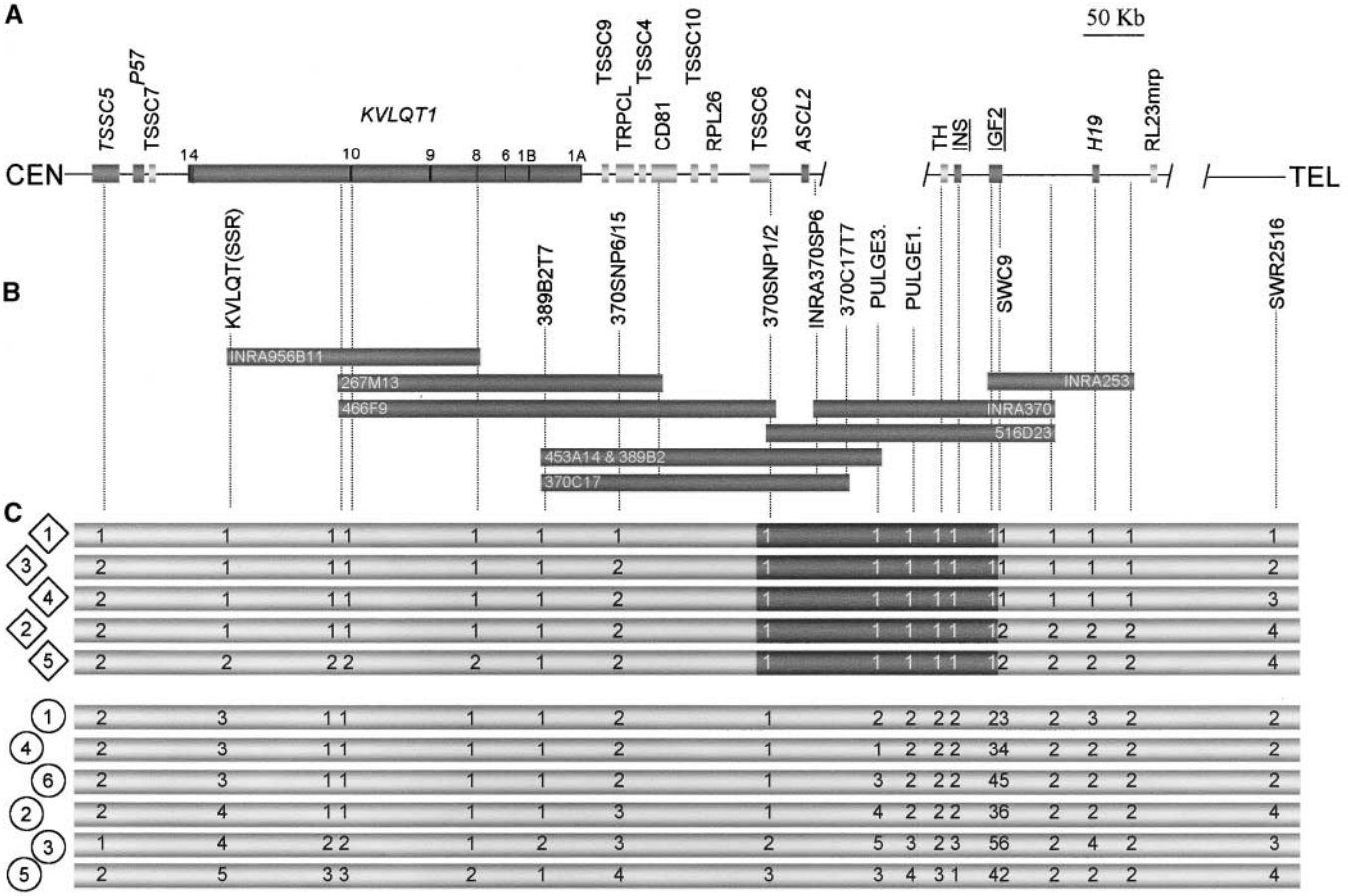

FIGURE 2.—(A) Schematic representation of the human 11p15 imprinted domain according to ONYANGO *et al.* (2000). Biallelically expressed genes are shown as lightly shaded cylinders, imprinted genes as darkly shaded cylinders. Maternally expressed genes are italicized and paternally expressed genes are underlined. (B) BAC contig spanning the porcine ortholog of the 11p15 imprinted domain, assembled by STS content mapping. The length of the horizontal bars does not reflect the actual physical size of the corresponding BACs. It is assumed in this map that the gene order with respect to telomere is conserved in man and pig (http://www.ensembl.org/homo_sapiens/). (C) Marker haplotypes of the five *Q* chromosomes (diamonds) and six *q* chromosomes (circles). Closely linked SNPs (<5 kb) or adjacent SNPs that could not be ordered were merged into polyallelic multisite haplotypes (*cf.* Table 2). The darkened chromosome segments correspond to the haplotype shared by all *Q* chromosomes and are therefore assumed to contain the QTL.

and were kept for further analysis (BACs 466F9, 267M13, 389B2, 370C17, and 453A14). To this set we added two BACs that were previously shown to span the *TH-H19* region (INRA370 and INRA253; AMARGER *et al.* 2002) and one that was isolated from the same library using a *KVLQT1* probe (INRA956B11; E. GIUFFRA, unpublished results). Three additional STS were developed from BAC end sequences (389B2T7, 370C17T7, and INRA370SP6) and 4 from randomly selected subclones of BAC 370C17 (370SNP1, 370SNP2, 370SNP6, and 370SNP15; Table 1). INRA370SP6 and 370SNP2 revealed highly significant BLAST hits in the vicinity of the *ASCL2* gene (expected values 10^{-7}) and *TSSC6* gene (expected values 10^{-18}), respectively, providing additional anchor points between the human and porcine sequence. Using STS content mapping, we assembled the BAC contig shown in Figure 2. It is compatible with the overall conservation of gene order between human and pigs in this chromosome region and indicates that the gap remaining in the human sequence between the *INS* and *ASCL2* genes may not be >55 kb.

All available STS were then amplified from genomic DNA of the 14 QTL-genotyped boars (see above) and cycle sequenced to identify DNA sequence polymorphisms. We identified a total of 51 single-nucleotide polymorphisms (SNPs): 2 in *TSSC5*, 15 in *KVLQT1*, 3 in *389B2-T7*, 1 in 370SNP15, 1 in 370SNP6, 3 in 370SNP1, 3 in 370SNP2, 4 in *TH*, 7 in *INS*, 4 in *IGF2*, 1 in *IG(IGF2-H19)*, 3 in *H19*, and 4 in IG(H19-RL23MRP) (Table 1).

Three microsatellites were added to this marker list: one (*KVLQT1(SSR)* isolated from BAC INRA956B11 and two (*PULGE1* and *PULGE3*) isolated from BAC INRA370 (Table 1).

Assembling pools of *Q- vs. q*-bearing chromosomes: To reconstruct the marker linkage phase of the 14 QTL-genotyped sires, we selected—for each boar—offspring that were homozygous for the alternate paternal *SWR-2516-SWC9* haplotypes. These were genotyped for all

TABLE 2

Definition of the multisite haplotypes (MHx) corresponding to the different markers shown in Figures 2 and 3

STS	MH1	MH2	MH3	MH4	MH5
TSSC5(I1)	T-T	C-C			
KVLQT1(I12)	C-C-C-G-C-C	C-T-T-A-T-C	T-C-C-A-T-T		
KVLQT1(I11)	T-C-G-C-T-T-G-T	G-T-A-T-C-C-A-T	G-C-G-C-C-C-G-G		
389B2T7	C-G-C	T-A-A			
370SNP6 + 370SNP15	C-(-)	T-(-)	C-C	T-C	
370SNP1 + 370SNP2	G-G-A-G-A-C	A-A-C-G-A-C	G-A-C-A-G-T		
TH1 + TH2	T-C-G-(CAAGGCCAGGT)	A-(-)-A-(-)	A-C-G-(CAAGGCCAGGT)		
INS(5′) + INS(I1,E2,I2) + INS(3′)	G-G-A-C-C-C-G	A-T-G-T-G-T-A	G-G-G-T-G-C-G		
IGF2(E3) + IGF2(E5) + IGF2(I8)	G-2-G-T	A-2-C-C	A-1-G-T	A-2-G-T	G-2-G-T
H19	C-C-T	C-C-C	C-T-C	T-T-C	
IG(H19-RL23mrp)A,B,C,D	C-A-T-T	T-G-C-C			

SNPs and microsatellites available in the region, and from these genotypes we manually determined the linkage phase of the boars.

For six of the seven boars, shown by marker-assisted segregation analysis to be of *Qq* genotype (Figure 1), the *Q* chromosomes associated with an increase in percentage of lean meat proved to be identical by state (IBS) over their entire length. This haplotype was therefore referred to as Q^1. The haplotype corresponding to the seventh *Q* chromosome (P7 in Figure 2) was different and was referred to as Q^2. For three of these sires, the haplotypes associated with a decrease in percentage of lean meat proved to be completely IBS as well and were thus referred to as q^1. The other four *q* chromosomes carried distinct haplotypes and were referred to as q^2, q^3, q^4, and q^5 (Figure 2).

The first boar that proved to be homozygous for the QTL by marker-assisted segregation analysis (P8) carried the q^4 haplotype on one of its chromosomes. Its other haplotype therefore had to be of *q* genotype as well and was referred to as q^6.

Boar P9 appeared to be heterozygous Q^1/Q^2. Boars P10 and P11 carried the Q^1 haplotype shared by six of the *Qq* boars. As a consequence, the other chromosomes of boars P10 and P11, which were IBS as well, were placed in the *Q* pool and referred to as Q^3. Homozygous boar P12 carried haplotype Q^2. As a consequence, its homolog was referred to as Q^4. Following the same recursive procedure, boars P13 and P14 were identified as being respectively Q^3Q^4 and Q^2Q^5.

The marker genotypes of the resulting five *Q* and five *q* chromosomes are shown in Figure 2. In this figure, closely linked (<5 kb) SNPs or SNPs that could not be ordered (370SNP15 with 370SNP6 and 370SNP1 with 370SNP2) were merged into a series of polyallelic multisite haplotypes. The correspondence between SNP genotype and haplotype number is given in Table 2.

All *Q* chromosomes share an ~250-kb common haplotype encompassing the *INS* and *IGF2* genes: Visual examination of the *Q* and *q* pools immediately reveals that all five chromosomes in the *Q* pool share an IBS haplotype spanning the *370SNP1/2-IGF2* interval as predicted (Figure 2). Further examination of the *Q* chromosomes indicates that on the proximal side of the *370SNP1/2-IGF2* interval, the Q^2, Q^3, and Q^4 haplotypes are all identical, while the Q^1 and Q^5 haplotypes are best explained by single-recombination events having occurred, respectively, in the *370SNP6/15-370SNP1/2* (Q^1) and *KVLQT1(I7-389B27)* (Q^5) interval. On the distal side of the *370SNP1/2-IGF2* interval, three of the five *Q* chromosomes (Q^1, Q^3, Q^4) carry the same haplotype in the *SWC9-IG(H19-RL23mrp)* interval, while the two remaining ones (Q^2, Q^5) are sharing a completely distinct one. Again, this is best explained by assuming a single ancestral recombination event just upstream of the *SWC9* microsatellite marker. These observations therefore strongly suggest that the hypothesized *Q* allele causing an increase in percentage of lean meat appeared by mutation or migration on a founder chromosome carrying the haplotype highlighted in Figure 2 and that the QTL is consequently located in the *370SNP6/15-SWC9* interval.

No such shared haplotype could be identified in the *q* pool. As expected under our model, the *q* pool exhibited a higher level of genetic diversity. The *q*-bearing chromosomes would indeed be older, having had ample opportunity to recombine, thereby increasing haplotype diversity. This can be quantified more accurately by computing the average pairwise probability for *Q* and *q* chromosomes to be IBD conditional on flanking marker data, using the coalescent model developed by Meuwissen and Goddard (2001). As shown in Figure 3, the average pairwise IBD probability among the five *Q* chromosomes is superior to 0.25 over the entire *KVLQT1(SSR)-IG(H19-RL23mrp)* interval and exceeds 0.60 in the *370SNP1/2-IGF2* interval. For the *q* chromosomes, the equivalent parameter is inferior to 0.20 over the entire studied region.

One could argue that the probability of identifying a shared haplotype among five chromosomes by chance alone is high and does not support the location of the

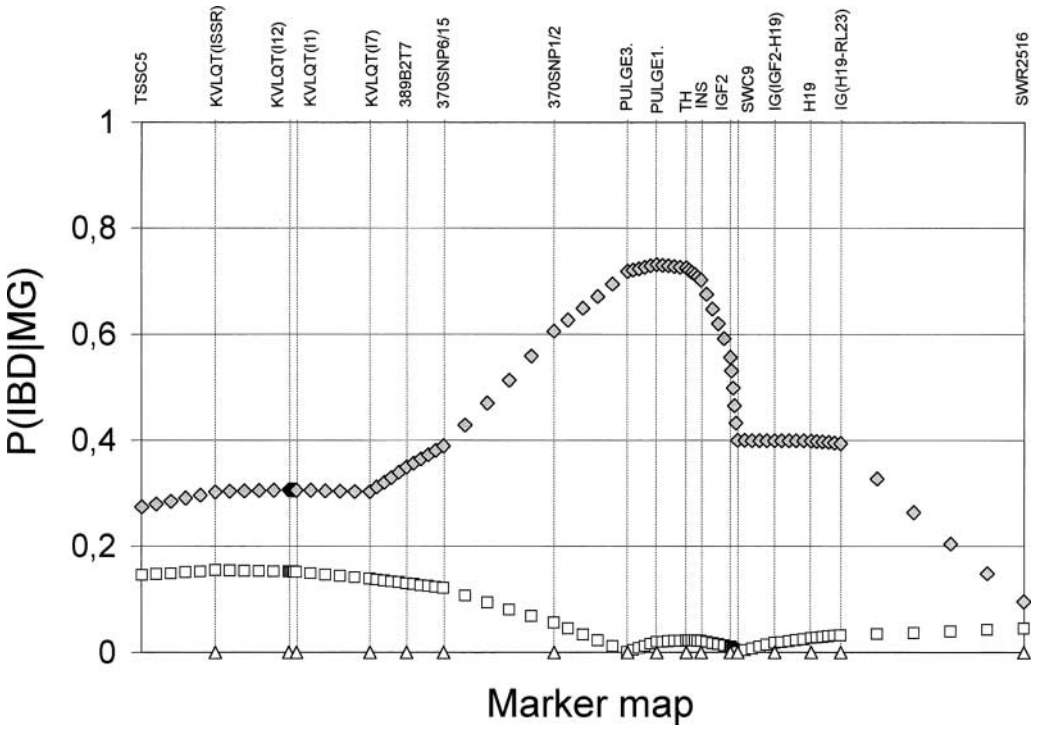

FIGURE 3.—Average probability for two Q (◆) and two q (□) chromosomes to be identical by descent at a given map position conditional on flanking marker data [P(IBD|MG)], along the chromosome segment encompassing the *p57-H19* imprinted domain, computed according to MEUWISSEN and GODDARD (2001). The positions of the markers defined according to Tables 1 and 2 are shown by the vertical dotted lines.

QTL within this region. To more quantitatively estimate the significance of the haplotype sharing among Q chromosomes, accounting for the distance between adjacent markers as well as allelic frequencies, we performed a multipoint LD analysis using the DISMULT program (TERWILLIGER 1995). To test the significance of the haplotype sharing observed among the five Q chromosomes, we performed the same DISMULT analysis on all 462 possible combinations of the 11 chromosomes taken 5 at a time. For each of these analyses we stored the highest likelihood obtained anywhere along the analyzed chromosome segment. The likelihood obtained using the five real Q chromosomes at the position of marker *PULGE3* was the highest one obtained across all chromosome permutations (data not shown), clearly indicating that the observed haplotype sharing is very unlikely to be fortuitous.

At present, our best estimate of the size of the *370SNP6/15-SWC9* interval is of the order of 250 kb. From the knowledge of the orthologous region in the human, it is predicted (ONYANGO *et al.* 2000) to contain a maximum of eight biallelically expressed genes (*TSSC9, TRPCL, TSSC4, CD81, TSSC10, RPL26, TSSC6,* and *TH*), two imprinted genes expressed from the maternal allele (*KVLQT1* and *ASCL2*), and two imprinted genes expressed from the paternal allele (*INS* and *IGF2*; Figure 2).

DISCUSSION

When we previously demonstrated that only the paternal SSC2 QTL allele influenced muscle mass and that the most likely QTL position coincided with *IGF2*, this gene stood out as a prime positional candidate (JEON *et al.* 1999; NEZER *et al.* 1999). On the basis of the initial linkage analysis, however, the confidence interval for the QTL covered ~4 cM, which were bound to contain a multitude of genes other than *IGF2*. It was therefore critical to refine the map position of the QTL, which we set out to do by exploiting both LD and marker-assisted segregation analysis. Because of the observed parent-of-origin effect, we focused our analysis on a chromosome region that is the ortholog of the human 11p15 imprinted domain. We herein provide strong evidence that the QTL indeed maps to the *p57-H19* imprinted gene cluster and within this region to a chromosome segment of ~250 kb containing *INS* and *IGF2* as the only known paternally expressed genes. The proximal breakpoint of the interval indeed coincides with the *SWC9* microsatellite marker known to be located in the 3′ UTR of *IGF2*. These findings therefore considerably strengthen the candidacy of *IGF2* in particular and justify its detailed analysis.

The fact that we succeeded in refining the map position of this QTL down to the subcentimorgan level supports its simple molecular architecture. Together with recent successes in positional cloning and identification of the mutations that underlie the major gene and QTL (*e.g.*, GROBET *et al.* 1997; MILAN *et al.* 2000; BLOTT *et al.* 2002; FREKING *et al.* 2002; GRISART *et al.* 2002; SMIT *et al.* 2003), this clearly indicates that at least part of the genetic variation of production traits in livestock is due to single mutations with large effects on the traits of interest.

The success of haplotype-sharing approaches in fine mapping QTL in livestock also suggests that QTL may be mapped in these populations by virtue of the haplotype signature resulting from intense selection on *Q* alleles, *i.e.*, haplotypes of unusual length given their population frequency. The feasibility of this approach has recently been demonstrated in human populations for two loci (*G6PD* and *CD40* ligand) known to be involved in resistance to malaria (Sabeti *et al.* 2002). Preliminary analyses suggest that similar haplotype signatures characterize the QTL described in this work as well as a positionally cloned QTL influencing milk yield and composition on bovine chromosome 14 (Grisart *et al.* 2002). QTL undergoing selection could thus be identified in livestock in the absence of phenotypic data.

We are grateful to Frédéric Farnir and Danny Wagenaar for their help with the ML and LD analysis. This work was supported by a grant from the Belgian "Ministère des Classes Moyennes et de l'Agriculture" and Gentec.

LITERATURE CITED

Amarger, V., M. Nguyen, A. S. Van Laere, C. Nezer, M. Georges *et al.*, 2002 Comparative sequence analysis of the INS-IGF2-H19 gene cluster in pigs. Mamm. Genome **13:** 388–398.

Andersson, L., 2001 Genetic dissection of phenotypic diversity in farm animals. Nat. Rev. Genet. **2:** 130–138.

Ardlie, K. G., L. Kruglyak and M. Seielstad, 2002 Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. **3:** 299–309.

Blott, S., J.-J. Kim, S. Moisio, A. Schmidt-Küntzel, A. Cornet *et al.*, 2003 Molecular dissection of a quantitative trait locus: a phenylalanine to tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics **163:** 253–266.

Darvasi, A., 1998 Experimental strategies for the genetic dissection of complex traits in animal models. Nat. Genet. **18:** 19–24.

Dunner, S., C. Charlier, F. Farnir, B. Brouwers, J. Canon *et al.*, 1997 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.

Fahrenkrug, S. C., G. A. Rohrer, B. A. Freking, T. P. Smith, K. Osoegawa *et al.*, 2001 A porcine BAC library with tenfold genome coverage: a resource for physical and genetic map integration. Mamm. Genome **12:** 472–474.

Flint, J., and R. Mott, 2001 Finding the molecular basis of quantitative traits: successes and pitfalls. Nat. Rev. Genet. **2:** 437–445.

Florini, J. R., D. Z. Ewton and S. A. Coolican, 1996 Growth hormone and the insulin-like growth factor system in myogenesis. Endocr. Rev. **17:** 481–517.

Freking, B. A., S. K. Murphy, A. A. Wylie, S. J. Rhodes, J. W. Keele *et al.*, 2002 Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. Genome Res. **12:** 1496–1506.

Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto *et al.*, 1995 Mapping quantitative trait loci controlling milk production by exploiting progeny testing. Genetics **139:** 907–920.

Grisart, B., W. Coppieters, F. Farnir, L. Karim, C. Ford *et al.*, 2002 Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT gene with major effect on milk yield and composition. Genome Res. **12:** 222–231.

Grobet, L., L. J. Royo Martin, D. Poncelet, D. Pirottin, B. Brouwers *et al.*, 1997 A deletion in the myostatin gene causes double-muscling in cattle. Nat. Genet. **17:** 71–74.

Hanset, R., C. Dasnois, S. Scalais, C. Michaux and L. Grobet, 1995 Effets de l'introgression dans le génome Piétrain de l'allèle normal au locus de sensibilité à l'halothane. Genet. Sel. Evol. **27:** 77–88.

Jeon, J. T., O. Carlborg, A. Törnsten, E. Giuffra, V. Amarger *et al.*, 1999 A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. Nat. Genet. **21:** 157–158.

Mackay, T. F. C., 2001 Quantitative trait loci in Drosophila. Nat. Rev. Genet. **2:** 11–20.

Mauricio, R., 2001 Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nat. Rev. Genet. **2:** 370–381.

Meuwissen, T. H., and M. E. Goddard, 2001 Prediction of identity by descent probabilities from marker-haplotypes. Genet. Sel. Evol. **33:** 605–634.

Milan, D., J. T. Jeon, C. Looft, V. Amarger, A. Robic *et al.*, 2000 A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. Science **288:** 1248–1251.

Nezer, C., L. Moreau, B. Brouwers, W. Coppieters, J. Detilleux *et al.*, 1999 An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. Nat. Genet. **21:** 155–156.

Nezer, C., L. Moreau, D. Wagenaar and M. Georges, 2002 Results of a whole genome scan targeting QTL for growth and carcass characteristics in a Piétrain × Large White intercross. Genet. Sel. Evol. **34:** 371–387.

Onyango, P., W. Miller, J. Lehoczky, C. T. Leung, B. Birren *et al.*, 2000 Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. Genome Res. **10:** 1697–1710.

Reik, W., and J. Walter, 2001 Genomic imprinting: parental influence on the genome. Nat. Rev. Genet. **2:** 21–32.

Riquet, J., W. Coppieters, N. Cambisano, J.-J. Arranz, P. Berzi *et al.*, 1999 Identity-by-descent fine-mapping of QTL in outbred populations: application to milk production in dairy cattle. Proc. Natl. Acad. Sci. USA **96:** 9252–9257.

Sabeti, P. C., D. E. Reich, J. M. Higgins, H. Z. Levine, D. J. Richter *et al.*, 2002 Detecting recent positive selection in the human genome from haplotype structure. Nature **419:** 832–837.

Smit, M., K. Segers, L. G. Carrascosa, T. Shay, F. Baraldi *et al.*, 2003 Mosaicism of solid gold supports the causality of a noncoding A-to-G transition in the determinism of the callipyge phenotype. Genetics **163:** 453–456.

Terwilliger, J. D., 1995 A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. Am. J. Hum. Genet. **56:** 777–787.

Communicating editor: C. Haley