

20. Lengyel, M., Szatmary, Z. & Erdi, P. Dynamically detuned oscillators account for the coupled rate and temporal code of place cell firing. *Hippocampus* **13**, 700–714 (2003).
21. Pike, F. G. *et al.* Distinct frequency preferences of different types of rat hippocampal neurones in response to oscillatory input currents. *J. Physiol. Lond.* **529**, 205–213 (2000).
22. Kamondi, A., Acsady, L., Wang, X. J. & Buzsaki, G. Theta oscillations in somata and dendrites of hippocampal pyramidal cells *in vivo*: Activity-dependent phase-precession of action potentials. *Hippocampus* **8**, 244–261 (1998).
23. Burgess, N., Recce, M. & O'Keefe, J. A model of hippocampal function. *Neural Netw.* **7**, 1065–1081 (1994).
24. Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J. & McNaughton, B. L. Population dynamics and theta rhythm phase precession of hippocampal place cell firing: A spiking neuron model. *Hippocampus* **6**, 271–280 (1996).
25. Jensen, O. & Lisman, J. E. Hippocampal CA3 region predicts memory sequences: Accounting for the phase precession of place cells. *Learn. Mem.* **3**, 279–287 (1996).
26. Wallenstein, G. V. & Hasselmo, M. E. GABAergic modulation of hippocampal population activity: Sequence learning, place field development, and the phase precession effect. *J. Neurophysiol.* **78**, 393–408 (1997).
27. Wood, E. R., Dudchenko, P. A. & Eichenbaum, H. The global record of memory in hippocampal neuronal activity. *Nature* **397**, 613–616 (1999).
28. Moita, M. A., Rosis, S., Zhou, Y., LeDoux, J. E. & Blair, H. T. Hippocampal place cells acquire location-specific responses to the conditioned stimulus during auditory fear conditioning. *Neuron* **37**, 485–497 (2003).
29. Kahana, M. J., Sekuler, R., Caplan, J. B., Kirschen, M. & Madsen, J. R. Human theta oscillations exhibit task dependence during virtual maze navigation. *Nature* **399**, 781–784 (1999).
30. O'Keefe, J. & Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature* **381**, 425–428 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank C. Lever, F. Cacucci, T. Wills, J. Ryan, D. Edwards and T. Hartley for technical assistance. This work was supported by the MRC and the Wellcome Trust. J.H. was a Rothermere Fellow.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.O'K. (j.okeefe@ucl.ac.uk).

A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig

Anne-Sophie Van Laere^{1*}, Minh Nguyen^{3*}, Martin Braunschweig¹, Carine Nezer³, Catherine Collette³, Laurence Moreau³, Alan L. Archibald⁴, Chris S. Haley¹, Nadine Buys⁵, Michael Tally⁶, Göran Andersson¹, Michel Georges³ & Leif Andersson^{1,2}

¹Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, and ²Department of Medical Biochemistry and Microbiology, Uppsala University, BMC, Box 597, SE-751 24 Uppsala, Sweden

³Department of Genetics, Faculty of Veterinary Medicine, University of Liege (B43), 20, bd. de Colonster, 4000 Liege, Belgium

⁴Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, Scotland, UK

⁵Gentec, Kapelbaan 15, 9255 Buggenhout, Belgium

⁶Tally Consulting, SE-11458 Stockholm, Sweden

* These authors contributed equally to this work

Most traits and disorders have a multifactorial background indicating that they are controlled by environmental factors as well as an unknown number of quantitative trait loci (QTLs)^{1,2}. The identification of mutations underlying QTLs is a challenge because each locus explains only a fraction of the phenotypic variation^{3,4}. A paternally expressed QTL affecting muscle growth, fat deposition and size of the heart in pigs maps to the *IGF2* (insulin-like growth factor 2) region^{5,6}. Here we show that this QTL is caused by a nucleotide substitution in intron 3 of *IGF2*. The mutation occurs in an evolutionarily conserved CpG island that is hypomethylated in skeletal muscle. The mutation

abrogates *in vitro* interaction with a nuclear factor, probably a repressor, and pigs inheriting the mutation from their sire have a threefold increase in *IGF2* messenger RNA expression in post-natal muscle. Our study establishes a causal relationship between a single-base-pair substitution in a non-coding region and a QTL effect. The result supports the long-held view that regulatory mutations are important for controlling phenotypic variation⁷.

The QTL affecting muscle growth, fat deposition and heart size was first identified in intercrosses between the European wild boar and Large White domestic pigs and between Piétrain and Large White pigs^{5,6}. The alleles from the Large White breed in the first cross and the Piétrain breed in the second cross increased muscle mass and reduced back-fat thickness. The QTL explained 15–30% of the phenotypic variation in muscle mass and 10–20% of the variation in back-fat thickness^{5,6}. We recently used a haplotype-sharing approach to refine the map position of the QTL⁸. We assumed that a new allele (*Q*) promoting muscle development occurred *g* generations ago on a chromosome carrying the wild-type allele (*q*). We also assumed that the favourable allele had gone through a selective sweep due to the strong selection for lean growth in commercial pig populations. A QTL genotype cannot be deduced directly from an individual's phenotype but the QTL genotype of sires can be determined by progeny testing and marker-assisted segregation analysis². Twenty-eight chromosomes with known QTL status were identified. All 19 *Q*-bearing chromosomes shared a haplotype in the 250-kilobase (kb) interval between the markers *370SNP6/15* and *SWC9* (*IGF2* 3' untranslated region), which was therefore predicted to contain the QTL. This region contains *INS* and *IGF2* as the only known paternally expressed genes. Given their known functions and especially the role of *IGF2* in myogenesis⁹, they stood out as prime positional candidates.

We re-sequenced one of the 19 *Q* chromosomes (P208) and six *q* chromosomes (each corresponding to a distinct marker haplotype) for a 28.6-kb segment containing *IGF2*, *INS* and the 3' end of *TH*. This chromosome collection was expanded by including *Q* and *q* chromosomes from the following: (1) a wild boar/Large White intercross segregating for the QTL⁵; (2) a Swedish Landrace boar showing no evidence for QTL segregation in a previous study¹⁰; and (3) *F*₁ sires from a Hampshire/Landrace cross and a Meishan/Large White intercross both showing no indication for QTL segregation (see Methods). The lack of evidence for QTL segregation shows that the boars are either homozygous *Q/Q* or *q/q*. A Japanese wild boar was included as a reference for the phylogenetic analysis and it was assumed to be homozygous wild type (*q/q*). We identified a total of 258 DNA sequence polymorphisms corresponding to one polymorphic nucleotide per 111 base pairs (bp) (Fig. 1). Two major and quite divergent clusters of haplotypes were revealed (Supplementary Fig. 1). The two established *Q* haplotypes from Piétrain and Large White animals (P208 and LW3) were identical to each other and to the chromosomes from the Landrace (LR) and Hampshire/Landrace (H205) animals, showing that the latter two must be of *Q* type as well. The absence of QTL segregation in the offspring of the *F*₁ Hampshire × Landrace boar carrying the H205 and H254 chromosomes implies that the latter recombinant chromosome is also of *Q* type. This places the causative mutation downstream from *IGF2* intron 1, in the region for which H254 is identical to the other *Q* chromosomes. The Large White chromosome (LW197) from the Meishan/Large White pedigree clearly clustered with *q* chromosomes, implying that the *F*₁ sire used for sequencing was homozygous *q/q* as no overall evidence for QTL segregation was observed in this intercross. This is consistent with a previous study showing that Meishan pigs carry an *IGF2* allele associated with low muscle mass¹¹. Surprisingly, the Meishan allele (M220) was nearly identical to the *Q* chromosomes but with one notable exception, it shared guanine with all *q* chromosomes at a position (*IGF2*-intron3-nucleotide 3072) where all *Q* chromosomes have adenine (Fig. 1).

Under a bi-allelic QTL model, the causative mutation would

correspond to a DNA polymorphism for which the two alleles segregate perfectly between *Q* and *q* chromosomes. The G to A transition at *IGF2*-intron3-3072 is the only polymorphism fulfilling this criterion, implying that it is the causative quantitative trait nucleotide (QTN)¹. We genotyped the founders and 12 F₁ boars for the putative QTN to verify the QTL status of animals from the Meishan/Large White intercross. All founders and F₁ boars were homozygous G/G (*q/q*) except one Large White founder and one F₁ boar, which were heterozygous A/G. A QTL analysis revealed that the heterozygous boar, but no other F₁ sire, showed clear evidence for segregation of a paternally expressed QTL, and the Meishan allele increased back-fat thickness as predicted (χ^2 with 1 degree of freedom = 7.75, $P = 0.005$). Including this, we have so far tested 13 large sire families where the sire is heterozygous A/G at the QTN, and all have shown evidence for QTL segregation. In contrast, we have tested more than 50 sires, representing several different breeds, genotyped as homozygous A/A or G/G without obtaining any evidence for the segregation of a paternally expressed QTL at the *IGF2* locus. The results provide conclusive genetic evidence that *IGF2*-intron3-G3072A is the causative mutation. The Meishan allele is apparently identical to the ancestral haplotype on which the mutation occurred.

IGF2-intron3-3072 is part of an evolutionarily conserved CpG island of unknown function¹² located between differentially methylated region 1 (DMR1) and a matrix attachment region previously

defined in mice¹³⁻¹⁵. The 94-bp sequence around the mutation shows about 85% identity to human, and the wild-type nucleotide at *IGF2*-intron3-3072 is conserved among eight mammalian species (Fig. 1). The QTN occurs 3 bp downstream of a conserved 8-bp palindrome. The methylation status of the 300-bp fragment centred on *IGF2*-intron3-3072 and containing 50 CpG dinucleotides was examined by bisulphite sequencing in four-month-old *Q*^{pat}/*q*^{mat} and *q*^{pat}/*Q*^{mat} pigs. The CpG island was methylated in liver (26% of CpGs methylated on average), whereas in skeletal muscle both paternal (pat) and maternal (mat) chromosomes were essentially unmethylated (including the *IGF2*-intron3-3071 C residue) irrespective of QTL genotype (3.4% of CpGs methylated on average, Fig. 2a; see also Supplementary Fig. 2).

To uncover a possible function for this element, we performed electrophoretic mobility shift assays (EMSA) using wild-type (*q*) and mutant (*Q*) sequences. Nuclear extracts were incubated with radioactively labelled *q* or *Q* double-stranded oligonucleotides. One specific complex (C1 in Fig. 2b) was obtained with the unmethylated wild-type (*q*) but not the mutant (*Q*) probe using extracts from murine C2C12 myoblasts. This complex was not obtained with the *q*^{*} probe, which has a methylated CpG at the QTN (Fig. 2b). A complex with approximately the same migration—but slightly weaker—was also detected in extracts from human HEK293 embryonic kidney cells and HepG2 hepatocytes. The specificity of the complex was confirmed as competition was obtained with 10-

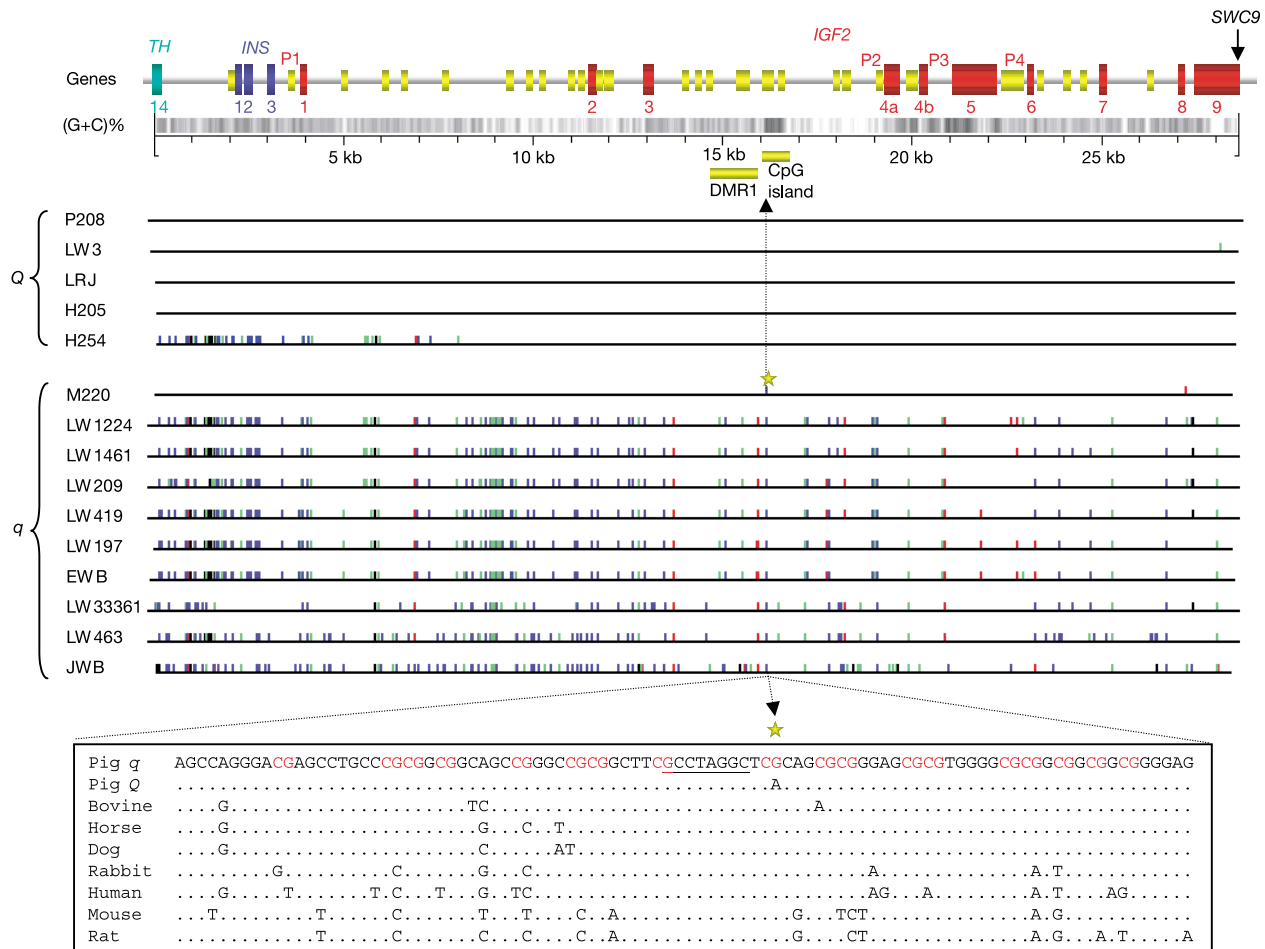


Figure 1 Polymorphisms in a 28.6-kb segment containing *TH* (exon 14), *INS* and *IGF2* among 15 pig chromosomes with deduced QTL status. The (G+C) content of a moving 100-bp window is shown on a grey scale (black 100%, white 0%). Yellow cylinders mark evolutionarily conserved regions¹². Viewgene²⁹ was used to highlight differences between the reference P208 sequence and other chromosomes. Blocks

indicate transitions (blue), transversions (green), insertions (red) and deletions (black). The QTN is indicated with an asterisk and the surrounding sequence is shown for eight mammals (CpGs are highlighted in red; a palindromic sequence is underlined). P, Piétrain; LW, Large White; LR, Landrace; H, Hampshire; M, Meishan; EWB, European wild boar; JWB, Japanese wild boar.

fold molar excess of unlabelled *q* probe, whereas a 50-fold excess of unlabelled *Q* probe or methylated *q** probe did not compete (Fig. 2b). Thus, the wild-type sequence binds a nuclear factor, and this interaction is abrogated by the mutation or methylation of the actual CpG site.

We analysed the effect of the *IGF2* *Q* mutation on transcription by using a transient transfection assay in mouse C2C12 myoblasts. We made *Q* and *q* constructs containing a 578-bp fragment from the actual region inserted in front of a luciferase reporter gene driven by the endogenous pig *IGF2* promoter 3 (P3) located approximately 5 kb downstream from the QTN (Fig. 1). This promoter was chosen because it generates the predominant *IGF2* transcript in muscle and the QTN affected the amount of P3 transcripts *in vivo* (see below). The *q* fragment clearly acted as a repressor element and reduced

luciferase activity to about 25%, whereas the *Q* fragment was a significantly weaker repressor element and showed about 70% activity compared with P3 alone (Fig. 2c). Our interpretation of this result, combined with those from the EMSA experiment, is that the *Q* mutation abrogates the interaction with a putative repressor protein. Thus, the *IGF2*-intron3-G3072A transition may be sufficient to explain the QTL effect as the two constructs only differ at this position. The constructs were also inserted in front of the heterologous herpes thymidine kinase (TK) minimal promoter. In this case the *q* construct caused a twofold increase of transcription whereas the *Q* construct caused a significantly higher, sevenfold increase (Supplementary Fig. 3). The results support our interpretation that the *q* allele represses transcriptional activity in C2C12 myoblasts when compared with the *Q* allele.

The *in vivo* effect of the mutation on *IGF2* expression was studied in a purpose-built *Q/q* × *Q/q* intercross. We tested the effect of the intron3-3072 mutation on *IGF2* imprinting, as a deletion encompassing DMR0, DMR1 and the associated CpG island derepresses the maternal *IGF2* allele in mesodermal tissues in mouse¹⁴. This was achieved by monitoring transcription from paternal and maternal alleles in tissues of *q/q*, *Q^{pat}/q^{mat}* and *q^{pat}/Q^{mat}* animals heterozygous for the *SWC9* microsatellite located in the *IGF2* 3' untranslated region. Before birth, *IGF2* was expressed exclusively from the paternal allele in skeletal muscle and kidney, irrespective of QTL genotype. At four months of age, weak expression from the maternal allele was observed in skeletal muscle, however at comparable rates for all three QTL genotypes (Supplementary Fig. 4). Only the paternal allele could be detected in four-month-old kidney. Consequently, the mutation does not seem to affect the

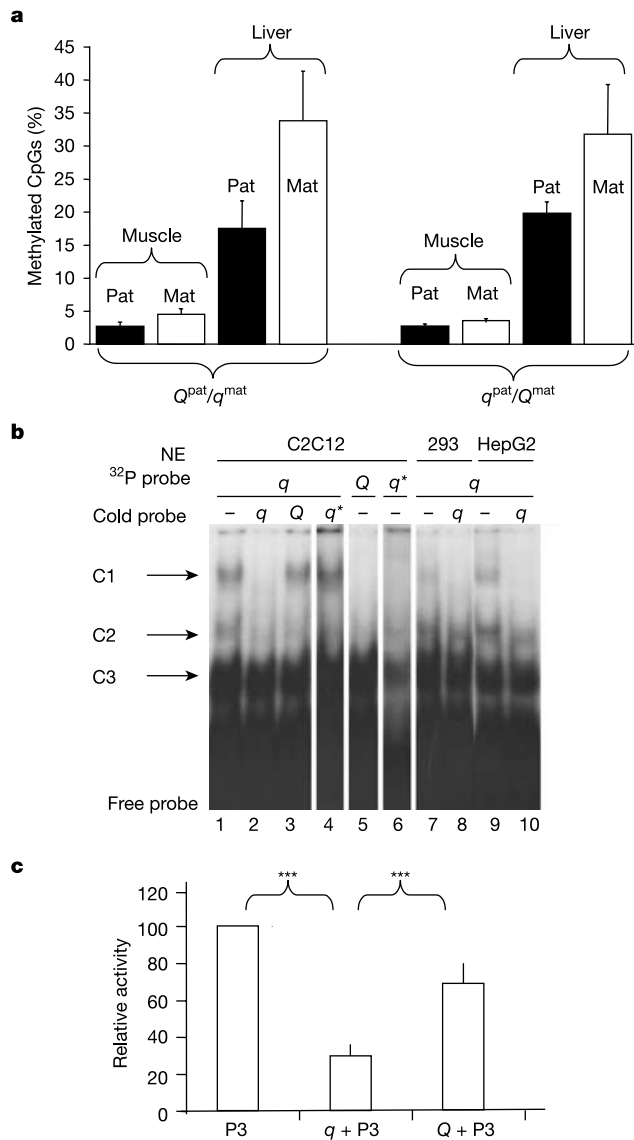


Figure 2 Methylation status, EMSA and transfection assays assessing the significance of the *IGF2* mutation. **a**, Percentage methylation around the QTN in liver and skeletal muscle of four-month-old pigs. Means ± s.e.m. are given; the numbers of analysed paternal (Pat) and maternal (Mat) chromosomes were in the range 10–26. **b**, EMSA using nuclear extracts (NE) from C2C12, HEK293 and HepG2 cells. Complex 1 (C1) was exclusively detected with the *q* probe. C2 was stronger in *q* but also probably present in the *Q* lane. C3 was unspecific. **c**, Luciferase assays of reporter constructs using pig *IGF2* P3 promoter and intron 3 fragments (*Q* and *q*). Relative activities compared with the P3–LUC reporter are reported (means ± s.e.m.). Triple asterisk, $P < 0.001$.

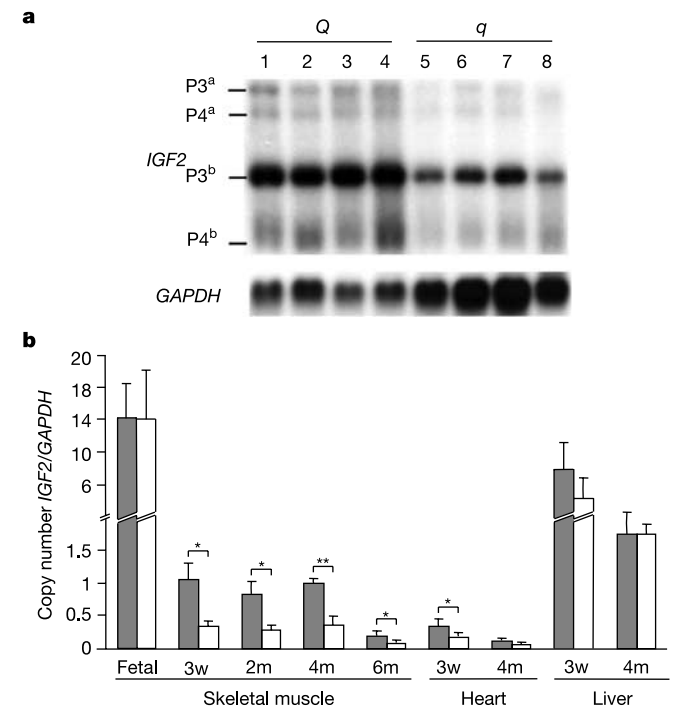


Figure 3 Analysis of *IGF2* mRNA expression. **a**, Northern blot of skeletal muscle (*gluteus*) poly(A)⁺ RNA from three-week-old piglets. Animals 1–4 and 5–8 carried a paternal *IGF2* *Q* or *q* allele, respectively. P3 and P4 indicate promoter usage, and a/b superscripts indicate the alternative polyadenylation signal used. All four transcripts showed a higher relative expression (standardized using *GAPDH*) in the *Q* group ($P < 0.05$). **b**, Real-time PCR analysis of *IGF2* expression in pigs carrying paternal *IGF2* *Q* (grey columns) or *q* (white columns) alleles. Expression levels were normalized using *GAPDH*. Means ± s.e.m. are given, $n = 3–11$. Asterisk, $P < 0.05$; double asterisk, $P < 0.01$; w, week; m, month.

imprinting status of *IGF2*. The partial derepression of the maternal allele in skeletal muscle may however explain why in a previous study muscle growth was found to be slightly superior in $q^{\text{pat}}/Q^{\text{mat}}$ versus q/q animals, and in Q/Q versus $Q^{\text{pat}}/q^{\text{mat}}$ animals⁵.

The *Q* allele was expected to be associated with increased *IGF2* expression because IGF-II stimulates myogenesis⁹. We monitored the relative mRNA expression of *IGF2* at different ages in the Q/q \times Q/q intercross using both northern blot analysis and real-time polymerase chain reaction (PCR) (Fig. 3). The expression levels in fetal muscle and postnatal liver at three weeks of age were approximately tenfold higher compared with postnatal muscle. No significant difference was observed in fetal samples or in postnatal liver samples, but a significant threefold increase of postnatal *IGF2* mRNA expression in skeletal muscle was observed in Q/Q or $Q^{\text{pat}}/q^{\text{mat}}$ versus $q^{\text{pat}}/Q^{\text{mat}}$ or q/q progeny. There was also a significant, but less pronounced, increase of mRNA expression in heart associated with the Q^{pat} allele. The significant difference in *IGF2* expression revealed by real-time PCR was confirmed using two different internal controls: GAPDH (Fig. 3b) and HPRT (data not shown). We found an increase of all detected transcripts originating from the three promoters (P2–P4) located downstream of the QTN. The results provide strong support for *IGF2* being the causative gene. The significant differences in *IGF2* mRNA expression between genotypes in skeletal and cardiac muscle and the lack of significant differences in fetal muscle and postnatal liver are consistent with our previous data showing clear phenotypic effects of the *IGF2* QTL on muscle growth and size of the heart but no effects on birth weight or weight of liver⁵. The results demonstrate that *IGF2* has an important role in regulating postnatal myogenesis. The higher expression in postnatal skeletal and cardiac muscle associated with the *Q* allele parallels the observation of a continued postnatal *IGF2* expression in mesodermal tissue of transgenic mice carrying a 5-kb deletion of *IGF2* encompassing DMR1 and the associated CpG island¹⁴.

Immunoreactive serum levels of IGF-II were determined by a radioimmunoassay, but no significant differences between genotypes were observed (Supplementary Fig. 5). This finding was expected because the major source of IGF-II in serum is generated from liver, where no difference in *IGF2* expression was detected (Fig. 3b). The results suggest that locally produced IGF-II in muscle determines the phenotype, as also indicated by the observation that there is no general overgrowth in pigs expressing the *Q* allele but rather a changed body composition.

There has been a strong selection for lean growth (high muscle mass and low fat content) in commercial pig populations over the past 50 yr. Therefore we investigated how this selection pressure has affected the allele frequency distribution of the *IGF2* QTL. The causative mutation was absent in a small sample of European and Asian wild boars and in breeds that have not been strongly selected for lean growth (Supplementary Table 1). In contrast, the causative mutation was found at high frequencies in several breeds that have been subjected to strong selection for lean growth. This confirms our prior assumption that *IGF2*^{*}*Q* has experienced a selective sweep and has been spread between breeds by cross-breeding. The *IGF2*^{*}*Q* mutation increases meat production, at the expense of fat, by 3–4%. The high frequency of *IGF2*^{*}*Q* among major pig breeds implies that this mutation has had a large impact on pig production. European and Asian pigs were domesticated from different subspecies of the wild boar, and Asian germplasm was introgressed into European pig breeds during the eighteenth and nineteenth centuries¹⁶. The *IGF2*^{*}*Q* mutation apparently occurred on an Asian chromosome as it shows a very close relationship to the haplotype carried by Chinese Meishan pigs. This explains the large genetic distance between *Q* and *q* haplotypes present in European domestic pigs (Fig. 1; see also Supplementary Fig. 1).

We have achieved an extraordinary resolution, down to a single nucleotide difference, in the genetic analysis of a QTL. This was

possible by exploiting a selective sweep of a favourable mutation in commercial pig populations. Determination of complete genome sequences from major farm animals in the near future will allow the exploitation of the full potential of farm animal genomics. High-density single-nucleotide polymorphism typing of domestic animals will allow identification of selective sweeps as documented here. Furthermore, re-sequencing of the entire genome using samples from different breeds, including wild ancestral species, will be very fruitful for studying genotype–phenotype relationships. This is well illustrated by the recent identification of the causative mutations for several interesting phenotypes in domestic animals^{17–23}. For instance, the *callipyge* mutation in sheep shares several features with the *IGF2* mutation in pigs; it affects body composition (muscle/fat content), shows a parent-of-origin effect and is a single-base substitution in a non-coding region²². In fact, farm animal genetics provides major advantages compared with human genetics, and in some aspects compared with model organisms, for genetic dissection of multifactorial traits². Farm animals are emerging as prime model organisms for understanding the genetic basis for multifactorial traits. □

Methods

Marker-assisted segregation analysis

QTL genotyping of the Pietrain/Large White, wild boar/Large White and Hampshire/Landrace crosses was performed as described⁸. Briefly, the likelihood of the pedigree data was computed under two hypotheses: H_0 , postulating that the corresponding boar was homozygous at the QTL (Q/Q or q/q), and H_1 , postulating that the boar was heterozygous Q/q . Likelihoods were computed using ‘percentage lean meat’ as phenotype, and assuming an allele substitution effect of 3.0%⁶. If the probability in favour of one of the hypotheses were superior or equal to 100:1, the most likely hypothesis was considered to be true. The Meishan/Large White cross consisted of 703 F_2 animals with data on back-fat depths. An interval analysis approach²⁴ with microsatellite markers spanning the *IGF2* region revealed no indication of an overall QTL effect, imprinted or not.

DNA sequencing

Animals homozygous for 13 of the haplotypes of interest were identified using flanking markers and pedigree information. A 28.6-kb segment was amplified from genomic DNA in seven long-range PCR products using the Expand Long Template PCR system (Roche Diagnostics GmbH). The same procedure was used to amplify the remaining M220 and LW197 haplotypes from two BAC clones isolated from a genomic library made from a Meishan/Large White F_1 individual²⁵. PCR products were purified using GeneClean (Polylab) and sequenced using the Big Dye Terminator Sequencing or dGTP Big Dye Terminator kits (Perkin Elmer). All primers are given in Supplementary Table 2. The sequence traces were assembled and analysed for DNA sequence polymorphism using Polyphred/Phrap/Consed²⁶.

Genotyping of *IGF2*-intron3-3072

Genotyping was done by pyrosequencing (Pyrosequencing AB). A 231-bp DNA fragment was PCR-amplified using Hot Star Taq DNA polymerase and Q-Solution (Qiagen) with the primers 18274F (5′-biotin-GGGCCGCGGCTTCGCTAG-3′) and 18274R (5′-CGCAGCTTCTCCTGCCACTG-3′). The sequencing primer (5′-CCCACGCGCTCCCGCGT-3′) was designed on the reverse strand because of the palindrome located 5′ of the QTN.

Bisulphite-based methylation analysis

Bisulphite sequencing was performed as described²⁷. A 300-bp fragment centred around intron3-3072 was amplified using a two-step PCR reaction with the following primers: PCR1-UP, 5′-TTGAGTGGGGATGTGAAGTTT-3′; PCR1-DN, 5′-ACCCACTTAT AATCTAAAAAATAATAATATATCTAA-3′; PCR2-UP, 5′-GGGGATTGTTGAAG TTTT-3′; PCR2-DN, 5′-CTTCTCCTACCCTAA-3′. The amplified strand was chosen in order to differentiate the *Q* and *q* alleles. PCR products were cloned in the pCR2.1 vector (Invitrogen). Plasmid DNA was purified (Plasmid mini kit, Qiagen) and sequenced (Big Dye Terminator kit, Perkin Elmer). Inserts with identical sequences, that is having the same combination of C residues (whether part of a CpG dinucleotide or not) converted to U, were considered to derive from the same PCR product and were only considered once.

Electrophoretic mobility shift assays

DNA-binding proteins were extracted as described²⁸. EMSAs were performed with 40 fmol ³²P-labelled, double-stranded oligonucleotide, 10 μ g nuclear extract and 2 μ g poly(dI-dC) in binding buffer (15 mM HEPES pH 7.65, 30.1 mM KCl, 2 mM MgCl₂, 2 mM spermidine, 0.1 mM EDTA, 0.63 mM dithiothreitol, 0.06% NP-40, 7.5% glycerol). For competition assays a 10-, 20-, 50- and 100-fold molar excess of cold double-stranded oligonucleotide were added. Reactions were incubated for 20 min on ice before ³²P-labelled, double-stranded oligonucleotide was added. Binding was allowed to proceed for 30 min at room temperature. DNA–protein complexes were resolved on a 5% native polyacrylamide gel

run in TBE $\times 0.5$ at room temperature for 2 h at 150 V. The following two (*Q/q*) 27-bp unmethylated oligonucleotides were used: 5'-GATCCTTCGCTAGGCTC(A/G)CAGCG CGGGAGCGA-3'. A methylated *q* probe (*q**) was generated by incorporating a methylated cytosine at the mutated CpG site during oligonucleotide synthesis.

Transient transfection assay

The constructs contained 578 bp from *IGF2* intron 3 (nucleotides 2868–3446), followed by the *IGF2* P3 promoter (nucleotides –222 to +45 relative to the start of transcription)¹² and a luciferase reporter. C2C12 myoblast cells were grown to approximately 80% confluence. Cells were transiently co-transfected with the firefly luciferase reporter construct (4 μ g) and a Renilla luciferase control vector (pHRG-TK, Promega; 80 ng) using 10 μ g Lipofectamine 2000 (Invitrogen). Cells were incubated for 25 h before lysis in 100 μ l Triton lysis solution. Luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega). The results are based on four triplicate experiments using two independent plasmid preparations for each construct. Statistical analysis was done with an analysis of variance.

Northern blot analysis and real-time RT-PCR

Total RNA was prepared using Trizol (Invitrogen) and treated with DNase I (Ambion). Products from the first-strand complementary DNA synthesis (Amersham Biosciences) were purified with QIAquick columns (Qiagen). Poly(A)⁺ RNA was then isolated using the Oligotex mRNA kit (Qiagen). Poly(A)⁺ mRNA (about 75 ng) from each sample was separated in a MOPS/formaldehyde agarose gel and transferred overnight to a Hybond-N⁺ nylon membrane (Amersham Biosciences). The membrane was hybridized in ExpressHyb hybridization solution (Clontech). The quantification of the transcripts was performed with a Phosphor Imager 425 (Molecular Dynamics). Real-time PCR was performed with an ABI PRISM 7700 instrument (Applied Biosystems). TaqMan probes and primers are given in Supplementary Table 3. PCRs were performed in triplicate using the Universal PCR Master Mix (Applied Biosystems). Messenger RNA was quantified using ten-point calibration curves established by dilution series of the cloned PCR products. Statistical evaluations were done with a two-sided Kruskal–Wallis rank-sum test.

Received 11 June; accepted 11 September 2003; doi:10.1038/nature02064.

- Mackay, T. F. C. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–21 (2001).
- Andersson, L. Genetic dissection of phenotypic diversity in farm animals. *Nature Rev. Genet.* **2**, 130–138 (2001).
- Glazier, A. M., Nadeau, J. H. & Aitman, T. J. Finding genes that underlie complex traits. *Science* **298**, 2345–2349 (2002).
- Darvasi, A. & Pisante-Shalom, A. Complexities in the genetic dissection of quantitative trait loci. *Trends Genet.* **18**, 489–491 (2002).
- Jeon, J.-T. *et al.* A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the *IGF2* locus. *Nature Genet.* **21**, 157–158 (1999).
- Nezer, C. *et al.* An imprinted QTL with major effect on muscle mass and fat deposition maps to the *IGF2* locus in pigs. *Nature Genet.* **21**, 155–156 (1999).
- King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
- Nezer, C. *et al.* Haplotype sharing refines the location of an imprinted QTL with major effect on muscle mass to a 250 Kb chromosome segment containing the porcine *IGF2* gene. *Genetics* **165**, 277–285 (2003).
- Florini, J. R., Ewton, D. Z. & McWade, F. J. IGFs, muscle growth, and myogenesis. *Diabetes Rev.* **3**, 73–92 (1995).
- Evans, G. J. *et al.* Identification of quantitative trait loci for production traits in commercial pig populations. *Genetics* **164**, 621–627 (2003).
- de Koning, D. J. *et al.* Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proc. Natl Acad. Sci. USA* **97**, 7947–7950 (2000).
- Amarger, V. *et al.* Comparative sequence analysis of the *Insulin-IGF2-H19* gene cluster in pigs. *Mamm. Genome* **13**, 388–398 (2002).
- Greally, J. M., Guinness, M. E., McGrath, J. & Zemel, S. Matrix-attachment regions in the mouse chromosome 7F imprinted domain. *Mamm. Genome* **8**, 805–810 (1997).
- Constancia, M. *et al.* Deletion of a silencer element in *Igf2* results in loss of imprinting independent of *H19*. *Nature Genet.* **26**, 203–206 (2000).
- Eden, S. *et al.* An upstream repressor element plays a role in *Igf2* imprinting. *EMBO J.* **20**, 3518–3525 (2001).
- Giuffra, E. *et al.* The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* **154**, 1785–1791 (2000).
- Grobet, L. *et al.* A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nature Genet.* **17**, 71–74 (1997).
- Milan, D. *et al.* A mutation in *PRKAG3* associated with excess glycogen content in pig skeletal muscle. *Science* **288**, 1248–1251 (2000).
- Galloway, S. M. *et al.* Mutations in an oocyte-derived growth factor gene (*BMP15*) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nature Genet.* **25**, 279–283 (2000).
- Mulsant, P. *et al.* Mutation in bone morphogenetic protein receptor-1B is associated with increased ovulation rate in Booroola Merino ewes. *Proc. Natl Acad. Sci. USA* **98**, 5104–5109 (2001).
- Pailhoux, E. *et al.* A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genet.* **29**, 453–458 (2001).
- Freking, B. A. *et al.* 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 (2002).
- Grisart, B. *et al.* Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition. *Genome Res.* **12**, 222–231 (2002).
- 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).

- Anderson, S. I., Lopez-Corralles, N. L., Gorick, B. & Archibald, A. L. A large fragment porcine genomic library resource in a BAC vector. *Mamm. Genome* **11**, 811–814 (2000).
- Nickerson, D., Tobe, V. O. & Taylor, S. L. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescent-based resequencing. *Nucleic Acids Res.* **25**, 2745–2751 (1997).
- Engemann, S., El-Maarri, O., Hajkova, P., Oswald, J. & Walter, J. in *Methods in Molecular Biology* Vol. 181 (ed. Ward, A.) (Humana Press, Totowa, New Jersey, 2002).
- Andrews, N. C. & Faller, D. V. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19**, 2499 (1991).
- Kashuk, C., Sengupta, S., Eichler, E. & Chakravarti, A. ViewGene: a graphical tool for polymorphism visualization and characterization. *Genome Res.* **12**, 333–338 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank C. Charlier and H. Ronne for discussions, M. Laita, B. McTeir, J. Pettersson, A.-C. Svensson and M. Köping-Höggård for technical assistance, and the Pig Improvement Company for providing DNA samples from Berkshire and Gloucester Old Spot pigs. This work was supported by the Belgian Ministère des Classes Moyennes et de l'Agriculture, the AgriFunGen program at the Swedish University of Agricultural Sciences, the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, Gentec, the UK Department for Environment, Food and Rural Affairs, the UK Pig Breeders Consortium, and the Biotechnology and Biological Sciences Research Council.

Competing interests statement The authors declare competing financial interests: details accompany the paper on www.nature.com/nature.

Correspondence and requests for materials should be addressed to L.A. (Leif.Andersson@imbim.uu.se) or M.G. (michel.georges@ulg.ac.be). The sequence data reported in this paper have been deposited in GenBank under accession numbers AY242098–AY242112.

Identification of the haematopoietic stem cell niche and control of the niche size

Ji Wang Zhang¹, Chao Niu¹, Ling Ye², Haiyang Huang², Xi He¹, Wei-Gang Tong¹, Jason Ross¹, Jeff Haug¹, Teri Johnson¹, Jian Q. Feng², Stephen Harris², Leanne M. Wiedemann³, Yuji Mishina³ & Linheng Li^{1,4}

¹Stowers Institute for Medical Research, Kansas City, Missouri 64110, USA
²Department of Oral Biology, School of Dentistry, University of Missouri–Kansas City, 650 East 25th Street, Kansas City, Missouri 64108, USA
³Laboratory of Reproductive and Developmental Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709, USA
⁴Department of Pathology and Laboratory Medicine, Kansas University Medical Center, Kansas City, Kansas 66160, USA

Haematopoietic stem cells (HSCs) are a subset of bone marrow cells that are capable of self-renewal and of forming all types of blood cells (multi-potential)¹. However, the HSC 'niche'—the *in vivo* regulatory microenvironment where HSCs reside—and the mechanisms involved in controlling the number of adult HSCs remain largely unknown. The bone morphogenetic protein (BMP) signal has an essential role in inducing haematopoietic tissue during embryogenesis^{2,3}. We investigated the roles of the BMP signalling pathway in regulating adult HSC development *in vivo* by analysing mutant mice with conditional inactivation of BMP receptor type IA (BMPRIA). Here we show that an increase in the number of spindle-shaped N-cadherin⁺ CD45[−] osteoblastic (SNO) cells correlates with an increase in the number of HSCs. The long-term HSCs are found attached to SNO cells. Two adherens junction molecules, N-cadherin and β -catenin, are asymmetrically localized between the SNO cells and the long-term HSCs. We conclude that SNO cells lining the bone surface function as a key component of the niche to support HSCs, and that BMP signalling through BMPRIA controls the number of HSCs by regulating niche size.