Genome-wide scan for body composition in pigs reveals important role of imprinting

Dirk-Jan de Koning*, Annemieke P. Rattink, Barbara Harlizius, Johan A. M. van Arendonk, E. W. Brascamp, and Martien A. M. Groenen

Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands

Communicated by James E. Womack, TexasA&M University, College Station, TX, May 12, 2000 (received for review December 22, 1999)

The role of imprinting in body composition was investigated in an experimental cross between Chinese Meishan pigs and commercial Dutch pigs. A whole-genome scan revealed significant evidence for five quantitative trait loci (QTL) affecting body composition, of which four were imprinted. Imprinting was tested with a statistical model that separated the expression of paternally and maternally inherited alleles. For back fat thickness, a paternally expressed QTL was found on *Sus scrofa* **chromosome 2 (SSC2), and a Mendelianexpressed QTL was found on SSC7. In the same region of SSC7, a maternally expressed QTL affecting muscle depth was found. Chromosome 6 harbored a maternally expressed QTL on the short arm and a paternally expressed QTL on the long arm, both affecting intramuscular fat content. The individual QTL explained from 2% up to 10% of the phenotypic variance. The known homologies to human and mouse did not reveal positional candidate genes. This study demonstrates that testing for imprinting should become a standard procedure to unravel the genetic control of multifactorial traits.**

It is well established that quantitative trait loci (QTL) under-
lying the genetic variance of multifactorial traits can be lying the genetic variance of multifactorial traits can be mapped in experimental as well as outbred populations (1, 2). Whole-genome scans have revealed a number of genomic regions contributing to genetic variation and have provided insight into the form of gene action. The genome scans can also be used to search for non-Mendelian forms of expression (3), but these opportunities have not been exploited systematically. Knowledge of mode of inheritance of identified QTL is important for medical and agricultural applications.

Parental genomes undergo modifications during gametogenesis, resulting, for some genes, in parent-of-origin-specific expression in the offspring. This phenomenon of genomic imprinting, as a form of epigenetic gene regulation, has been shown to influence several subchromosomal areas in mammals (4). In human and mouse, most imprinted genes are arranged in chromosomal clusters†, and their linked organization suggests coordinated mechanisms controlling imprinting and gene expression (5, 6). It is generally viewed that imprinting is involved in fetal growth and brain development (7).

Different approaches have been used over time to identify imprinted areas in the genome. Both Robertsonian and reciprocal translocations resulting in mice with uniparental disomy for portions of the genome have been used to identify imprinted regions on six chromosomes (8). Furthermore, chromosomal anomalies associated with imprinted diseases in humans helped to identify imprinted genes and to narrow regions of interest (9, 10). More recently, molecular genetic approaches taking advantage of, for example, methylation patterns observed for imprinted genes, have been used to isolate imprinted genes (11–14). The number of known genes is increasing rapidly, but imprinting has been reported only for about 30 (8). In livestock, evidence for imprinting was found for one specific chromosomal region in sheep and one in pigs (15–17). Imprinting effects, however, have not been studied systematically for multifactorial traits. We present results of a genome-wide approach to detect imprinted regions for multifactorial traits in an experimental cross of pigs.

Materials and Methods

Experimental Population. Boars from the Chinese Meishan pig breed were crossed with sows from commercial Dutch pig lines. From the resulting F_1 , randomly selected boars and sows were mated to create the F_2 population (18). This experimental population facilitates the dissection of the genetics underlying phenotypic differences between these breeds for body composition traits. Meishan pigs are characterized by high fatness compared with Dutch pigs, which have been selected for lean growth for many generations. On 785 F_2 pigs, we recorded three body composition traits after slaughter: back fat thickness and muscle depth measured between the third and fourth rib, and percentage of intramuscular fat inside the *Musculus longissimus* (18). The phenotypic mean (\pm SD) of the F₂ population was 22.0 (± 5.7) mm for back fat thickness, 40.6 (± 6.7) mm for muscle depth, and 1.84 (\pm 0.87)% for intramuscular fat content (18). Assuming Mendelian expression, analyses for back fat thickness and intramuscular fat content on part of this population revealed significant evidence for QTL on chromosome 2 and on chromosome 7 affecting back fat thickness (19).

Genotyping and Statistical Analyses. A whole-genome scan including a test for imprinting was used to map autosomal QTL on the F2 population. Genotypes were obtained for 132 microsatellite markers, covering more than 90% of the porcine genome, which were selected after testing many markers on the individual Meishan grandfathers and DNA pools of the grandmother lines (19). Genotypes were obtained for the F_2 animals, their F_1 parents, and the purebred Meishan grandparents.

The statistical analyses were based on the line cross concept (20), where original breeds are assumed homozygous for different QTL alleles but can have marker alleles in common. Extension of this model to test for imprinting has been suggested (3) and used in the analysis of the *IGF2* region in pigs (17). Analysis with this model, however, provided evidence for imprinting, but a separate test was needed to infer paternal or maternal expression. The model for imprinting (3), therefore was reparameterized to enable a direct test for the contribution of the paternally and maternally inherited effect. For every F_2 individual, we inferred the probabilities of inheriting two Meishan alleles (P_{11}) , two Dutch alleles (P_{22}) , or one from each line $(P_{12}$ or P_{21} ,

Abbreviations: QTL, quantitative trait loci; cM, centimorgan; SSC, *Sus scrofa* chromosome.

^{*}To whom reprint requests should be addressed. E-mail: Dirk-Jan.deKoning@alg.vf.wau.nl.

[†]A world wide web site is provided by C. V. Beechey, B. M. Cattanach, and R. L. Selley (Medical Research Council, Mammalian Genetics Unit, Harwell, Oxfordshire, U.K.): http:// www.mgu.har.mrc.ac.uk/imprinting/imptables.html.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: Proc. Natl. Acad. Sci. USA, 10.1073/pnas.140216397. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.140216397

Table 1. Genetic model for QTL affecting three body composition traits

*Partial *F* ratio for the individual components of a model including a paternal, maternal, and dominance component at the most likely position of the QTL. †Empirical confidence intervals obtained by bootstrapping for the relevant model.

‡Estimates of QTL effects for the inferred genetic model. The additive effect (Mendelian expression) and the paternal or maternal effect (imprinting) are expressed as the deviation of the Meishan allele (3, 19, 20). Standard errors of the estimates are in parentheses.

 ${}^{5}P$ < 0.0001.

different subscripts according to parental origin; first subscript is paternally inherited allele) at 1-centimorgan (cM) intervals across the genome. Using multiple marker information for a given location in the genome, we calculated the probability of the two alleles in an offspring corresponding to any of the four possible combinations (3, 20). The probabilities are functions of the recombination rates between the location under consideration and the flanking informative markers, which may vary from progeny to progeny depending on the genotype of the F_1 parents and the Meishan grandparents. Under the traditional line cross approach, an additive effect (*a*) and a dominance effect (*d*) are estimated by using the regression of the phenotypes on $P_{a} = P_{11} - P_{22}$ and $P_{d} = P_{12} + P_{21}$. To separate the contribution of the parents, we introduced the probability that the individual inherited a Meishan allele from its father $(P_{\text{pat}} = [P_{11} + P_{12}]$ $[P_{22} + P_{21}]$ or from its mother $(P_{\text{mat}} = [P_{11} + P_{21}] - [P_{22} + P_{12}]).$ A saturated model, which included a paternal (P_{pat}) , a maternal (P_{mat}) , and a dominance component (P_d) , was fitted at 1-cM intervals across the genome. For each position of a QTL, the mode of inheritance of the QTL was inferred based on the contribution of each of the three components. The contribution of a component was measured by the reduction in total sum of squares caused by incorporating that component in the model after fitting the other components. The *F* statistic was used to evaluate the significance of each component. This evaluation facilitated discrimination between QTL showing exclusive paternal expression, exclusive maternal expression, or Mendelian expression.

Significance Thresholds and Confidence Intervals. For the inferred genetic models, the significance thresholds and the confidence intervals of the QTL position were determined empirically. The significance threshold was set at the 5% genome-wise risk level (21). This threshold accounted for testing the entire genome but not for testing multiple traits. These thresholds were determined by permutation with at least 10,000 replicates (19).

Empirical confidence intervals for the QTL position were obtained by bootstrapping the data followed by analysis of the replicates under the inferred genetic model. From each of 10,000 bootstrap replicates, the best test statistic was stored. The 95% cutoff point of the sorted (in descending order) test statistics provided an empirical threshold to define the boundaries of the confidence interval. This method is an alternative to other bootstrapping strategies in which QTL positions of the replicates are sorted to determine an empirical confidence interval (3). The method used here allows for noncontinuous confidence intervals and is closer to the traditional logarithm of odds drop-off methods.

Results

Our genome scan resulted in five significant QTL affecting body composition traits, of which four were imprinted. For back fat thickness, there was strong evidence for a paternally expressed QTL on *Sus scrofa* chromosome 2 (SSC2; Table 1). For the QTL affecting back fat thickness on SSC7, both the paternal and maternal component were highly significant, implying Mendelian expression for this QTL. For muscle depth, a highly significant QTL mapped to the same area as the QTL for back fat thickness on SSC7. In contrast to the QTL for back fat thickness, the QTL for muscle depth was maternally expressed (Table 1). From these results, it cannot be determined whether there are two linked loci or one locus with pleiotropic effects that shows imprinting during one stage of development and Mendelian expression during another.

With a model ignoring imprinting, suggestive evidence for a Mendelian QTL for intramuscular fat content was reported on the long arm of SSC6 (19). The present analysis, however, revealed that this effect was caused by a significant paternally expressed QTL (Table 1). In addition, a maternally expressed QTL affecting the same trait was found on the short arm of the same chromosome. The phenotypic variance explained by the individual QTL varied from 2% for the QTL affecting intramuscular fat content on SSC6 to 10% for the QTL affecting back fat thickness on SSC7.

A graphical comparison of results obtained under the imprinting and Mendelian models is shown in Fig. 1. The imprinted QTL for back fat thickness on SSC2 maps 35 cM from the *IGF2* region, for which an imprinted QTL for muscularity and fat deposition has been reported (16, 17). Although the confidence interval does not exclude *IGF2* as a candidate gene, our results indicate that an additional imprinted QTL is present more proximal on this chromosome. The reported QTL in the *IGF2* region primarily controlled muscularity (16, 17), whereas in the present study, we found no evidence for a QTL affecting muscle depth on SSC2. All three studies provided convincing evidence for a QTL, which rules out chance as a cause for the observed differences in affected traits between studies. The discrepancies, however, might very well be due to the differences in founder populations, in particular between the Piétrain, wild boar, and Meishan breeds. Also, differences in age and weight at which carcass composition was measured may play a role.

The general outline of the comparative map between pig and human for the regions of interest has been established[‡] by using bidirectional chromosome painting, a somatic cell hybrid panel,

 $*$ The comparative map of the pig can be viewed at http://www.toulouse.inra.fr/lgc/pig/ cyto/cyto.htm. Alignment of the porcine cytogenetic and linkage map is adapted from http://sol.marc.usda.gov/genome/swine/htmls/chromosome_list.html.

Fig. 1. Test statistic profiles for three porcine chromosomes that exhibit imprinting effects for one of the body composition traits: SSC2 and back fat thickness (*A*), SSC6 and intramuscular fat content (*B*), SSC7 and muscle depth (*C*), and SSC7 and back fat thickness (*D*). The black line represents the test statistic for a Mendelian QTL vs. no QTL. The blue line represents the test statistic for a paternally expressed QTL vs. no QTL. The red line represents the test statistic for a maternally expressed QTL vs. no QTL. The black horizontal line denotes the 5% genome-wise threshold for the Mendelian model, and the blue line indicates the same threshold for the imprinting models (thresholds for maternal and paternal expression were very similar and well within the sampling variance associated with permutation testing). Homologous regions in humans are indicated as colored bars (22–24, 26).‡ Imprinted genes located within these human chromosomal areas are listed at the bottom (5, 25).

and fluorescent *in situ* hybridization (refs. 22–24; Fig. 1). Genes that have been mapped more precisely in pigs, by linkage analysis or on the radiation hybrid panel (26), facilitated further refinement of the comparative map. We realize that the comparative map presented herein is not comprehensive and that some genes originating from other chromosomes are reported but not represented in Fig. 1.

QTL affecting body composition traits in pigs can have implications for obesity research in humans (20). Although several obesity-related disorders that are reported in humans and mice map to homologous regions of the imprinted QTL found in this study (27), imprinting has been reported only for the Prader–Willi Syndrome (HSA15q11.2-q12; refs. 9 and 10).

The QTL on SSC7 can be narrowed to a region homologous with HSA6p21.3-p22. This region contains the major histocompatibility complex, including *LTA*, and shows extensive conservation in gene order (28). Imprinted genes have not been reported for this region in humans or mice (5).

For the maternally expressed QTL affecting intramuscular fat content on SSC6p, several genes that map to the area are located on HSA16q22-ter. No imprinted genes have been reported for this region in humans. For the paternally expressed QTL affecting intramuscular fat content on SSC6q, candidate genes *MC5R* (29), *FABP3* (30), and *UOX* (26) map between markers SW316 and S0003. These genes are located on human chromosomes 18p11.2, 1p33-p32, and 1p22, respectively, and in humans, imprinting has not been reported for these regions. However, the confidence interval of this QTL extends on both sides to homologous regions in humans, where imprinted genes have been reported: *p73* on HSA1p36 and *PEG3* on HSA19q13.4 (imprinted only in mice).

For SSC2, imprinting is reported for the *IGF2* area, but until now, homology to other imprinting clusters could not be established clearly. Data on imprinting of the Wilms' Tumor gene 1 (*WT1*) on HSA11p13 are contradictory (5).

Discussion

The progress of the genome projects, in particular the large number of polymorphisms that have been characterized in many species, has boosted the search for genes involved in multifactorial traits such as obesity, diabetes, and schizophrenia. Genomic imprinting, however, is regarded to be a rare phenomenon and consequently is ignored in most studies. Our results indicate that genomic imprinting might be a more common phenomenon than previously thought. We detected five QTL, of which four were subject to imprinting. For at least two of these regions, imprinting has not been reported in pigs, and the known homologies to humans and mice did not reveal obvious positional imprinted candidates. To our knowledge, only one study has considered imprinting in a genome-wide analysis, and these results indicated that uniparental expression, both paternal and maternal, might indeed be involved in diabetes (31).

The statistical analysis presented herein provides information on the mode of expression of genes. In addition, analysis under different modes of expression increases the power of finding genes. This increase is exemplified by the results for intramuscular fat content on SSC6, where significant evidence for QTL was found only under the imprinting model. The approach is implemented in this study for a cross between outbred lines but can be extended to other designs and methods of analysis, including mapping methods used in human genetic studies. For implementation of the method proposed herein, it is essential that parental origin of marker alleles can be derived for the offspring. This requirement excludes studies based on F_2 crosses or a single backcross between inbred lines that are commonly used in mice and rats (3). These model species have contributed enormously to the current understanding of genetic variation. The inability to detect imprinting in the most commonly used mapping designs has certainly contributed to the current feeling that imprinting is a rare phenomenon. The problem can be overcome by producing one backcross population from F_1 fathers as well as one from F_1 mothers as applied by Clapcott *et al.* (32) to demonstrate genomic imprinting for a major QTL controlling susceptibility to trypanosomiasis in mice. Outbred crosses, such as the cross between two pig breeds in our study, are the ideal resource for detection of imprinted regions.

The model of analysis assumes that alleles at the QTL are fixed in the parental lines. The QTL will be detected when the parental lines carry different alleles, which is likely given the marked morphological divergence between European and Chinese Meishan pigs. If the fixation assumption is violated and the alleles at the QTL are still segregating in either or both of the lines, the power of its detection will be greatly reduced, and its effect will be underestimated (33). Extreme QTL allele frequency differences between male and female parents could lead to the false identification of imprinting for a Mendelian QTL. In our study, this risk is small, because male and female parents were selected

- 1. Lander, E. S. & Schork, N. J. (1994) *Science* **265,** 2037–2048.
-
- 2. Paterson, A. H. (1995) *Genome Res.* **5,** 321–333. 3. Knott, S. A., Marklund, L., Haley, C. S., Andersson, K., Davies, W., Ellegren, H., Fredholm, M., Hansson, I., Hoyheim, B., Lundstrom, K., *et al*. (1998) *Genetics* **149,** 1069–1080.
- 4. Cattanach, B. M. & Beechey, C. V. (1997) in *Genomic Imprinting in the Mouse: Possible Final Analysis*, eds. Reik, W. & Surani, A. (IRL, Oxford), pp. 118–145.
- 5. Morison, I. M. & Reeve, A. E. (1998) *Hum. Mol. Genet.* **7,** 1599–1609.
- 6. Constancia, M., Pickard, B., Kelsey, G. & Reik, W. (1998) *Genome Res.* **8,** 881–900. 7. Tilghman, S. M. (1999) *Cell* **96,** 185–193.
- 8. Beechey, C. V. (1999) in *Genomic Imprinting: An Interdisciplinary Approach*, ed. Ohlsson, R.
- (Springer, Berlin), pp. 303–313. 9. Nicholls, R. D., Saitoh, S. & Horsthemke, B. (1998) *Trends Genet.* **14,** 194–200.
- 10. Ohta, T., Gray, T. A., Rogan, P. K., Buiting, K., Gabriel, J. M., Saitoh, S., Muralidhar, B., Krajewska-Walasek, M., Driscoll, D. J., Horsthemke, B., *et al*. (1999) *Am. J. Hum. Genet.* **64,** 397–413.
- 11. Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S. C., Ishino, F. & Surani, M. A. (1995) *Nat. Genet.* **11,** 52–59.
- 12. Hagiwara, Y., Hirai, M., Nishiyama, K., Kanazawa, I., Ueda, T., Sakaki, Y. & Ito, T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 9249–9254.
- 13. Gabriel, J. M., Higins, M. J., Gebuhr, T. C., Shows, T. B., Saitoh, S. & Nicholls, R. D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14857–14862.
- 14. Peters, J., Wroe, S. F., Wells, C. A., Miller, H. J., Bodle, D., Beechey, C. V., Williamson,
- C. M. & Kelsey, G. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 3830–3835. Cockett, N. E., Jackson, S. P., Shay, T. L., Farnir, F., Berghmans, S., Snowder, G. D., Nielsen, D. M. & Georges, M. (1996) *Science* 273, 236–238.
- 16. Nezer, C., Moreau, L., Brouwers, B., Coppieters, W., Detilleux, J., Hanset, R., Karim, L.,
Kvasz, A., Leroy, P. & Georges, M. (1999) *Nat. Genet. 21*, 155–156.
17. Jeon, J.-T., Carlborg, O., Tornsten, A., Giuffra, E.,
-

randomly from the same F_1 population. Furthermore, a large number of parents reduces the chance of allele frequency differences caused by sampling.

Genome-wide screens for QTL often result in estimates of QTL positions that lack precision, which complicates the identification of the responsible gene. Knowledge of the fact that the QTL is subject to imprinting will help in identifying the genes. Expression studies aimed at the identification of monoallelic expression of positional candidates will further aid the identification of the gene(s) responsible for the observed QTL effect. Genotypes of the parents can be used to discriminate between random inactivation and parent-of-origin effects.

For the practice of animal breeding, identification of major imprinted loci affecting body composition has several implications. Our results call for a revision of methods for genetic evaluation that currently ignore non-Mendelian expression. The net result of gametic imprinting is a reduction of the expected phenotypic covariance between parents and offspring relative to that between siblings. Identification of imprinted loci opens new perspectives for crossbreeding, which is common practice in pig breeding. Imprinted genes could further accommodate differentiation between sow lines, which are required to have optimal body composition to support their reproductive performance, and between boar lines, to ensure high-quality pork.

Although the mechanisms underlying imprinting are not totally unraveled (5), this study clearly demonstrates the important role of imprinting for body composition traits. We strongly urge, therefore, the inclusion of statistical testing for imprinting in human and animal genetic research, both in genome scans and in evaluating candidate genes.

The authors acknowledge technical assistance of P. van Oers, B. de Vries, T. Lenferink, M. Faivre, R. Acar, R. Joosten, and P. de Groot. This research was supported financially by the Netherlands Technology Foundation (STW) and was coordinated by the Earth and Life Sciences Foundation (ALW). Additional financial support was provided by the Dutch Product Board for Livestock, Meat, and Eggs and the Dutch pig breeding organizations Hypor BV, Dumeco Breeding BV, and Topigs. The European Union provided financial support for B.H. We acknowledge the U.S. Department of Agriculture-supported U.S. Pig Genome Coordination Project for contributed primers.

Eklund, L., Andersson, K., Hansson, I., Lundstrom, K., *et al*. (1999) *Nat. Genet.* **21,** 157–158.

- 18. Janss, L. L. G., Van Arendonk, J. A. M. & Brascamp, E. W. (1997) *Genetics* **145,** 395–408.
19. de Koning, D.-J., Janss, L. L. G., Rattink, A. P., Van Oers, P. A. M., De Vries, B. J., Groenen, M. A. M., Van Der Poel, J. J., De Groot, P. N., Brascamp, E. W. & Van Arendonk, J. A. M.
- (1999) Genetics 152, 1679–1690.
20. Andersson, L., Haley, C. S., Ellegren, H., Knott, S. A., Johansson, M., Andersson, K.,
Andersson-Eklund, L., Edfors-Lilja, I., Fredholm, M., Hansson, I., et al. (1994) *Science* 263, 1771–1774.
- 21. Lander, E. S. & Kruglyak, L. (1995) *Nat. Genet.* **11,** 241–247.
- 22. Goureau, A., Yerle, M., Schmitz, A., Riquet, J., Milan, D., Pinton, P., Frelat, G. & Gellin, J. (1996) *Genomics* **36,** 252–292.
- 23. Rettenberger, G., Klett, C., Zechner, U., Kunz, J., Volgel, W. & Hameister, H. (1995) *Genomics* **26,** 372–378.
- 24. Yerle, M., Lahbib-Mansais, Y., Mellink, C., Gourreau, A., Pinton, P., Echard, G., Gellin, J., Zijlstra, C., De Haan, N., Bosma, A. A., *et al*. (1995) *Mamm. Genome* **9,** 176–186.
- 25. Pearsall, R. S., Imai, K., Shibata, H., Hayashizaki, Y., Chapman, V. M., Held, W. A. & Plass,
- C. (1998) Mamm. Genome 9, 261–262.
26. Hawken, R. J., Murtaugh, J., Flickinger, G. H., Yerle, M., Robic, A., Milan, D., Gellin, J.,
Beattie, C. W., Shook, L. B. & Alexander, L. J. (1999) Mamm. Genome 10, 824–830.
- 27. Changnon, Y. C., Perusse, L., Weisnagel, J., Rankinen, T. & Bouchard, C. (2000) *Obes. Res.* **8,** 89–117.
- 28. Chardon, P., Renard, C. & Vaiman, M. (1999) *Immunol. Rev.* **167,** 179–192.
- 29. Kim, K. S., Marklund, S. & Rothschild, M. F. (2000) *Anim. Genet.*, **31,** 229–239. 30. Gerbens, F., Rettenberger, G., Lenstra, J. A., Veerkamp, J. H. & Te Pas, M. F. (1997) *Mamm. Genome* **8,** 328–332.
- 31. Paterson, A. D., Naimark, D. M. J. & Petronis, A. (1999) *Hum. Hered.* **49,** 197–204.
- 32. Clapcott, S. J., Teale, A. J. & Kemp, S. J. (2000) *Parasite Immunol. (Oxf.)* **22,** 259–264. 33. Alfonso, L. & Haley, C. S. (1998) *Anim. Sci.* **66,** 1–8.
-