doi:10.1111/j.1365-2052.2008.01763.x

Genetic variation in the *mannosidase 2B2* gene and its association with ovulation rate in pigs

E. M. Campbell[1], D. J. Nonneman, L. A. Kuehn and G. A. Rohrer

USDA, ARS, US Meat Animal Research Center, Clay Center, NE 68933-0166, USA

Summary

Ovulation rate is an important phenotypic trait that is a critical component of litter size in pigs. Despite being moderately heritable in pigs, selection for increased ovulation rate is difficult because it is difficult to measure and is a sex-limited trait. A QTL for ovulation rate residing on the p-terminal end of pig chromosome 8 has been detected in a Meishan-cross resource population. Comparative analysis of this region yielded a positional candidate gene, *mannosidase 2B2* (*MAN2B2*), for this QTL. The entire coding region of *MAN2B2* was resequenced in the Meishan and White Composite founder animals of the resource population to identify SNPs. Eleven polymorphisms that alter the protein product of *MAN2B2* were discovered and tested for statistical associations with ovulation rate in three generations of the resource population. The polymorphism located at position 1574 of the mRNA (D28521:c.1574A>G) was the most significant polymorphism tested ($P = 0.00005$) where the additive effect of the c.1574A allele was estimated to be −0.89 ova. This polymorphism was determined to be more significantly associated with ovulation rate than the breed-specific analysis conducted during the line-cross QTL discovery. The c.1574A>G marker was not associated with ovulation rate in an occidental population. Therefore, either *MAN2B2* has a unique epistatic interaction within the Meishan-cross population or the c.1574A>G SNP is in linkage disequilibrium with the actual causative genetic variation in the Meishan-cross population.

Keywords association, MAN2B2, ovulation rate, pigs, SNP.

Introduction

The use of genetic markers associated with the component traits of litter size in pigs could dramatically improve reproductive efficiency. One of the more heritable components of litter size is ovulation rate (Bennett & Leymaster 1989; Lamberson 1990; Young *et al.* 1996). A QTL for ovulation rate was identified on the terminal end of the p-arm of porcine chromosome 8 (Rohrer *et al.* 1999) and was subsequently verified in a larger group of Meishan–White Composite animals (Campbell *et al.* 2003). The refined maximum likelihood location of the QTL resided directly over the gene *mannosidase 2B2* (*MAN2B2*), which was physically assigned to this region by Ohata *et al.* (1997) and linkage-mapped by Campbell *et al.* (2003). The gene *MAN2B2* encodes an extracellular, ubiquitously expressed α-D-mannosidase (Jin *et al.* 1999; Tascou *et al.* 2000). Mannose subunits are removed from glycosylated proteins by α-D-mannosidases. The presence of oligosaccharide moieties on follicle-stimulating hormone (FSH), lutenizing hormone (LH) and thyroid-stimulating hormone is believed to modify activity and/or clearance rates of these hormones (Baenziger 1996). Because FSH is an important regulator of oocyte recruitment and LH is involved in ovulation, factors that affect glycosylation of these hormones could significantly affect ovulation rate. Therefore, *MAN2B2* was sequenced in Meishan and White Composite germplasm to determine if genetic variation in the coding region was associated with variation in ovulation rate in the USMARC Meishan–White Composite resource population.

Address for correspondence

G. A. Rohrer, USDA, ARS, US Meat Animal Research Center, Spur 18D, PO Box 166, Clay Center, NE 68933-0166, USA.
E-mail: gary.rohrer@ars.usda.gov

[1]Present address: 22339 E. Calle deFlores, Queen Creek, AZ 85242, USA.

Accepted for publication 27 April 2008

Materials and methods

Sequencing and SNP identification

Primers were developed to amplify a 219-bp region from the 3′ end of *MAN2B2* based on the cDNA sequence (D28521).

The forward primer sequence was 5′-GTAGCTCCAGACG TCCCTTG-3′ and the reverse primer was 5′-TTTAAGCA GTCCGAACACCC-3′. The amplicon was labelled with [α-^{32}P] dATP and used to screen the RPCI-44 porcine BAC library (BACPAC Resources) as described (Fahrenkrug *et al.* 2001). Five positive BAC clones were selected and grown, and BAC DNA was extracted using Qiagen plasmid midi preps. To facilitate sequencing, RPCI-44 BAC clone 355 I24 was subcloned into cosmids. Two overlapping cosmids that contained the entire gene were identified and grown in 500 ml of LB with ampicillin, and DNA was extracted using the Qiagen midi prep very low copy protocol.

Methods used to obtain the sequence of the intron–exon boundaries included direct sequencing from cosmid clones or sequencing plasmid subclones. Reactions for direct sequencing consisted of 4 μl template, 4 μl Big Dye (Perkin Elmer) and 6.4 pmol primer. Additionally, cosmids were subcloned by digesting with EcoRI (Promega), Tsp509I (New England BioLabs), BamHI (Promega), Sau3A (Promega) and PstI (Promega) and ligated into the pBluescript SK vector. Random and selected colonies were picked from plates and grown in 5 ml LB/ampicillin overnight. Plasmid DNA was then extracted using QIAprep® spin miniprep kits (Qiagen) and sequenced using M13 primers. Each sequencing reaction contained 3.2 pmol primers, 2 μl template and 2 μl Big Dye.

Primers were designed within introns to amplify entire exons (Table S1). Each exon was amplified in two Meishan and two White Composite founder boars with the most progeny in the resource population. The amplicons were purified using QIAquick PCR purification kits (Qiagen) and QIAquick gel extraction kits (Qiagen), and then sequenced. Each sequencing reaction contained 2 μl purified amplicon, 2 μl Big Dye and 3.2 pmol primer. If polymorphisms were detected in the first four boars sequenced, the remaining 16 grandparents were sequenced for that exon. When a polymorphism was detected that altered the amino acid composition, an assay was designed for the Sequenom MassArray genotyping platform using the sME chemistry as described previously (Nonneman & Rohrer 2003). Primers used for genotyping are listed in Table S2.

SNP associations with ovulation rate

The phenotypic data used for this study were described by Rohrer *et al.* (1999). Briefly, 10 Meishan founders (five males, five females) were mated to 10 White Composite founders (five female, five male; ¼ Chester White, ¼ Landrace, ¼ Large White, ¼ Yorkshire) to develop a resource population for genomic studies. Six hundred animals were phenotyped for ovulation rate in the backcross through F_4 generations of this resource population. Ovulation rates in backcross and F_4 generation gilts were determined by laparoscopic observation on pregnant unilaterally hysterectomized and ovariectomized gilts; ovulation rates on F_3 gilts were collected after slaughter at approximately 10 months of age.

Statistical analyses were as described by Kuehn *et al.* (2007). Regressors for additive and dominant effects of each SNP were calculated using genotypic probabilities computed by allelic peeling algorithms in an extended version of GENOPROB (Thallman 2002). The association model included fixed effects of contemporary group and SNP regressors, and random additive polygenic effects to account for genetic similarities between animals not explained by known genotypes. Variance components for polygenic effects and error were estimated using MTDFREML (Boldman *et al.* 1995). Initially, each SNP was fitted separately due to potential multicollinearity between regressions of closely linked markers. All significant SNPs were included in a common model and individually removed in successive iterations based on significance to form a final model. In addition to SNP genotypes, a fixed effect of breed of origin for the *MAN2B2* region was determined with GENOPROB. This variable was used to determine if any SNP markers were more significant than the breed-specific association previously reported (Campbell *et al.* 2003).

A composite population of animals (½ Landrace, ¼ Duroc and ¼ Yorkshire) with ovulation rate data (n = 680) was genotyped for the c.156T>C, c.1574A>G and c.1998C>A polymorphisms (D28521). Single-marker association analyses were conducted as described above and included all available pedigree information to estimate polygenic effects. Data were from three consecutive generations of animals. The frequency of the c.156T allele was too low to analyse in this population.

Results and discussion

The total size of the published cDNA sequence for *MAN2B2* (D28521) is 3313 bp. Sequence data from this study were deposited into GenBank as accession number EU346203. This study determined that *MAN2B2* has 19 exons by comparing sequence information derived from genomic clones with the cDNA sequence reported (D28521). The exon sizes are presented in Table S1. The exon sizes for both exon 1 and exon 19 include the untranslated regions. Twenty-six polymorphisms were detected in the transcript of *MAN2B2*. Twenty-one polymorphisms resided in the protein-coding region (Table 1), with 11 changing the amino acid composition of the mature protein. No polymorphisms were detected in exons 3, 5, 6, 7, 14, 16, 17 or 18.

In exon 1, there were four polymorphisms, three of which were in the 5′ untranslated region. One (c.156T>C) was in the putative start codon in the published sequence. It is located in the second base of the start codon and would inactivate this translational start site. A second start codon

Table 1 Polymorphisms detected in *MAN2B2*.

Exon	Location[1]	Polymorphism	Effect on protein
1	109	A>G	5′ UTR
1	122	G>A	5′ UTR
1	145	A>G	5′ UTR
1	156	T>C	Start codon
2	391	G>A	None
4	734	A>G	Thr>Ala
4	746	A>G	Thr>Ala
8	1297	G>C	None
8	1303	C>T	None
8	1330	G>A	Met>Ile
8	1344	A>G	His>Arg
8	1360	G>C	None
9	1414	G>A	None
10	1574	A>G	Met>Val
11	1723	T>C	None
11	1770	A>G	Lys>Arg
11	1897	C>T	None
11	1898	A>G	Ser>Gly
12	1998	C>A	Ala>Glu
13	2217	C>T	Ala>Val
13	2344	C>T	None
13	2392	C>T	None
15	2600	G>T	Val>Leu
19	3118	C>T	None
19	3216	T>C	3′ UTR
19	3217	T>C	3′ UTR

[1]The first nucleotide corresponds to the first position in the original published sequence (D28521).

Table 2 Single-marker association analyses of marker genotypes with ovulation rate.

SNP	Meishan freq.[1]	White Composite freq.[1]	a[1]	d[2]	P-value[3]
c.156T>C	0.05	0.05	−0.44	NF	0.20835
c.734A>G	0.10	0.10	−0.72	NF	0.04943
c.746A>G	0.70	0.10	−0.52	−0.03	0.03238
c.1330G>A	0.80	1.00	0.49	−0.36	0.13067
c.1344A>G	1.00	0.10	−0.84	0.08	0.00017
c.1574A>G	1.00	0.25	−0.89	0.07	0.00005
c.1770A>G	0.05	0.05	−0.59	NF	0.09397
c.1898A>G	0.10	0.20	−0.46	0.11	0.41744
c.1998C>A	1.00	0.20	−0.75	−0.17	0.00147
c.2217C>T	0.70	1.00	0.53	0.16	0.11979
c.2600G>T	0.40	1.00	0.25	0.29	0.22073

[1]Values for allelic frequencies and additive effects (a) are presented in terms of the first nucleotide in the SNP code, which is the nucleotide from the original published sequence (D28521).
[2]d, dominance effect. When the minor allele frequency was extremely low, the dominance effect was not fitted (NF) due to colinearity between additive and dominance regression coefficients.
[3]The P-value is the overall genotypic P-value of the F-ratio, which had one numerator degree of freedom when dominance was not fitted or two numerator degrees of freedom when both additive and dominance were fitted.

was present in the same reading frame 33 bases downstream from the published start site. The allele frequencies in the grandparents for the c.156T>C polymorphism (Table 2) indicate that the reported start site was rare in our population. Two founder animals were heterozygous for c.156T>C, whereas the rest of the animals were homozygous for the newly discovered c.156C allele. One of the heterozygous grandparents at c.156T>C was a Meishan boar, whereas the other was a White Composite sow. This allelic distribution was not expected for a causative polymorphism of this QTL as it was detected by comparing the average Meishan allele to the average White Composite allele. The causative polymorphism would be expected to have considerably different allelic frequencies between the founder breeds. As expected, the association analyses did not find a significant association between the c.156T>C polymorphism and ovulation rate (Table 2).

Most of the changes in amino acids would not be considered to have large effects on the mature protein because the substituted amino acids had similar biochemical properties. Four polymorphisms exchanged non-polar amino acids (c.1330G>A, c.1574A>G, c.2217C>T and c.2600G>T; methionine with isoleucine, methionine with valine, alanine with valine and valine with leucine respectively), two polymorphisms substituted positively charged amino acids (c.1344A>G and c.1770A>G; histidine with arginine and lysine with arginine respectively) and a final polymorphism exchanged uncharged polar amino acids (c.1898A>G; serine with glycine). Both polymorphisms in exon 4 (c.734A>G and c.746A>G) switched the non-polar amino acid alanine with the uncharged polar amino acid threonine. However, the c.1998C>A polymorphism exchanged alanine, a non-polar amino acid, with glutamic acid, a negatively charged amino acid.

There were no breed-specific SNPs among the 20 founder animals in the population (Table 2). Three SNPs were nearly breed-specific where the published allele was fixed in the Meishan founders while that allele was infrequent in the White Composite founders (c.1344A, 0.10; c.1998C, 0.20; and c.1574A, 0.25). For the c.2600G>T polymorphism, the c.2600T allele discovered in this study was not present in the White Composite while it had a frequency of 0.60 in Meishan founder pigs. For the remaining amino acid-altering polymorphisms, the most frequent allele was the same in both breeds and they had similar allelic frequencies (Table 2).

The results of the association analyses for each individual protein-altering SNP are presented in Table 2. As should be expected, the associations of markers with extremely different allelic frequencies between the founding breeds

(c.746A>G, c.1344A>G, c.1574A>G and c.1998C>A) and ovulation rate were highly significant. The only other nominally significant association was detected for c.734A>G. However, this association would not be significant if an adjustment for multiple comparisons was made. The most significant association detected was with c.1574A>G despite its greater deviation from breed specificity than c.1344A>G or c.1998C>A. To determine the most significant marker or set of markers, a stepwise model reduction was conducted with the four nominally significant markers included in the initial model. The only marker remaining in the final model was c.1574A>G. The estimated additive effect of c.1574A>G was −0.89 ova per copy of c.1574A ($P = 0.00002$) while the dominance effect was not significantly different from zero.

We then tested if c.1574A>G was more significantly associated with ovulation rate than a breed-specific analysis for this genomic location (i.e. what a line-cross QTL model would identify). Breed-specific regression coefficients were derived with GENOPROB. When both breed-specific coefficients and coefficients for c.1574A>G were fitted simultaneously, the only significant effect was the additive effect of c.1574A>G ($P = 0.043$). Although the significance level of c.1574A>G was similar to the maximum *F*-ratio found in Campbell *et al.* (2003), the estimated effect was considerably less in this study. This is likely due to fitting a polygenic effect because the breed-specific regression coefficients also had lower estimated effects than observed by Campbell *et al.* (2003).

Because the c.1574A>G polymorphism was the most significantly associated SNP with ovulation rate, all other significant associations detected were likely due to linkage disequilibrium with the c.1574A>G marker and/or the causative genetic variation. Unfortunately, not enough is known about the *MAN2B2* gene to determine if the c.1574A>G polymorphism could potentially alter the enzymatic activity of the mature protein. This marker needed to be tested in other populations to determine if the association was similar to the observed effect in this population, which would support the hypothesis that the c.1574A>G polymorphism may be the causative variation.

Therefore, an association analysis was conducted in a Landrace–Duroc–Yorkshire composite population with 680 ovulation rate records. There was no evidence that the c.1574A>G polymorphism was associated with variation in ovulation rate ($P = 0.698$). Likewise, the c.1998C>A polymorphism was not associated with ovulation rate in this population. Because the Landrace–Duroc–Yorkshire composite population did not have similar associations for c.1574A>G, the association reported in the Meishan-cross population is likely due to linkage disequilibrium or is a result of epistatic interactions unique to this research population. Additional research needs to be conducted to determine the genetic cause of this region's effect on ovulation rate.

Acknowledgements

The authors thank Kris Simmerman for genotyping and Sherry Kluver for manuscript preparation.

References

Baenziger J.U. (1996) Glycosylation: to what end for the glycoprotein hormones? *Endocrinology* **137,** 1520–2.

Bennett G.L. & Leymaster K.A. (1989) Integration of ovulation rate, potential embryonic viability and uterine capacity into a model of litter size in swine. *Journal of Animal Science* **67,** 1230–41.

Boldman K.G., Kriese L.A., Van Vleck L.D., Van Tassel C.P. & Kachman S.D. (1995) *A Manual for the Use of* MTDFREML. *A Set of Programs to Obtain Estimates of Variance and Covariances (Draft)*. USDA, Agriculture Research Service, Clay Center, NE.

Campbell E.M.G., Nonneman D. & Rohrer G.A. (2003) Fine mapping a quantitative trait locus affecting ovulation rate in swine on chromosome 8. *Journal of Animal Science* **81,** 1706–14.

Fahrenkrug S.C., Rohrer G.A., Freking B.A., Smith T.P.L., Osoegawa K., Shu C.L., Catanese J.J. & de Jong P.J. (2001) A porcine BAC library with tenfold genome coverage: a resource for physical and genetic map integration. *Mammalian Genome* **12,** 472–4.

Jin Y.Z., Dacheux F., Dacheux J.L., Bannai S., Sugita Y. & Okamura N. (1999) Purification and properties of major α-D-mannosidase in the luminal fluid of porcine epididymis. *Biochimica et Biophysica Acta* **1432,** 382–92.

Kuehn L.A., Rohrer G.A., Nonneman D.J., Thallman R.M. & Leymaster K.A. (2007) Detection of single nucleotide polymorphisms associated with ultrasonic backfat depth in a segregating Meishan × White Composite population. *Journal of Animal Science* **85,** 1111–9.

Lamberson W.R. (1990) Genetic parameters for reproductive traits. In: *Genetics of the Pig, North Central Regional Research Projects NC-103 Report* (Ed. by L.D. Young), pp. 70–6. USMARC, Clay Center, NE and University of Nebraska, Lincoln, NE.

Nonneman D.J. & Rohrer G.A. (2003) Comparative mapping of a region on chromosome 10 containing QTL for reproduction in swine. *Animal Genetics* **34,** 42–6.

Ohata K., Okamura N., Kojima M. & Yasue H. (1997) Assignment of *alpha-mannosidase* gene (*MAN2B2*) to swine chromosome 8p23-pter by fluorescence *in situ* hybridization. *Mammalian Genome* **8,** 158–9.

Rohrer G.A., Ford J.J., Wise T.H., Vallet J.L. & Christenson R.K. (1999) Identification of quantitative trait loci affecting female reproductive traits in a multigenerational Meishan-White Composite swine population. *Journal of Animal Science* **77,** 1385–91.

Tascou S., Nayernia K., Engel W. & Burfeind P. (2000) Refinement of the expression pattern of a mouse homologue of the porcine 135 kDa α-D-*mannosidase* (*MAN2B2*). *Biochemical and Biophysical Research Communications* **272,** 951–2.

Thallman R.M. (2002) *User's Manual for* GENOPROB, *Version 2.000*. USDA-ARS, US Meat Animal Research Center, Clay Center, NE.

Young L.D., Leymaster K.A. & Christenson R.K. (1996) Opportunities for indirect selection for uterine capacity of swine. *Journal of Animal Science* **74**(Suppl. 1), 119.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Primers used to amplify exons.

Table S2 Primers used for genotyping assays.

Please note: Blackwell Publishing is not responsible for the content or functionality of any supporting information supplied by the authors.