BMC Genetics

Genome-wide association and identification of candidate genes for age at puberty in swine

Dan J. Nonneman[*], James F. Schneider, Clay A. Lents, Ralph T. Wiedmann, Jeffrey L. Vallet and Gary A. Rohrer

Abstract

Background: Reproductive efficiency has a great impact on the economic success of pork production. Gilts comprise a significant portion of breeding females and gilts that reach puberty earlier tend to stay in the herd longer and be more productive. About 10 to 30 % of gilts never farrow a litter and the most common reasons for removal are anestrus and failure to conceive. Puberty in pigs is usually defined as the female's first estrus in the presence of boar stimulation. Genetic markers associated with age at puberty will allow for selection on age at puberty and traits correlated with sow lifetime productivity.

Results: Gilts ($n = 759$) with estrus detection measurements ranging from 140–240 days were genotyped using the Illumina PorcineSNP60 BeadChip and SNP were tested for significant effects with a Bayesian approach using GENSEL software. Of the available 8111 five-marker windows, 27 were found to be statistically significant with a comparison-wise error of $P < 0.01$. Ten QTL were highly significant at $P < 0.005$ level. Two QTL, one on SSC12 at 15 Mb and the other on SSC7 at 75 Mb, explained 16.87 % of the total genetic variance. The most compelling candidate genes in these two regions included the growth hormone gene (*GH1*) on SSC12 and *PRKD1* on SSC7. Several loci confirmed associations previously identified for age at puberty in the pig and loci for age at menarche in humans.

Conclusions: Several of the loci identified in this study have a physiological role for the onset of puberty and a genetic basis for sexual maturation in humans. Understanding the genes involved in regulation of the onset of puberty would allow for the improvement of reproductive efficiency in swine. Because age at puberty is a predictive factor for sow longevity and lifetime productivity, but not routinely measured or selected for in commercial herds, it would be beneficial to be able to use genomic or marker-assisted selection to improve these traits.

Keywords: Bayes, Genome-wide association, Puberty, Pig, SNP, Swine

Background

Puberty is the process of physical maturation of an animal to be capable of sexual reproduction. In humans puberty may be defined as age at first menses in girls, a milestone well-recalled and widely recorded. In pigs puberty is defined as age at first estrus when an animal will stand for breeding. With these definitions, puberty manifests as a spontaneous event. In reality, attainment of puberty is a complex maturation process that involves multiple body tissues and organ systems [1–3]. The hypothalamus receives neural and endocrine input from these systems to appropriately activate the pituitary-ovarian axis under conditions favorable for successful pregnancy to occur. Timing of puberty varies widely within and between populations and can be associated with several adult conditions and phenotypes. Increased risk for adult ovarian cancer [4–6], endometrial cancer [7, 8] and obesity [9–11] have been associated with early puberty in girls. Girls that reach puberty later have lower fertility [12, 13]. Similarly in pigs, later age at puberty is associated with lower fertility. Gilts that have an earlier age at first estrus stay in the herd longer, are more likely

* Correspondence: Dan.Nonneman@ars.usda.gov
United States Department of Agriculture, Agricultural Research Service, U.S.
Meat Animal Research Center, Clay Center, NE 68933, USA

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 2 of 9

to farrow multiple litters, give birth to more piglets and thus have a longer more productive life [14, 15].

Gilts comprise a significant portion of breeding females and thus successful gilt development is critical to overall herd performance. Management decisions prior to first mating of gilts can affect productivity and later reproductive performance [16, 17]; however, estrous traits of gilts (e.g., duration of estrus) are genetically correlated to adult reproductive phenotypes such as wean to estrus interval and sow longevity [18], making age at first estrus an early indicator trait that can be used to select for favorable adult reproductive performance. Age at puberty is moderately heritable in pigs ($h^2 = 0.38$ to 0.46) [19, 20]; but, age at first estrus is rarely recorded in pork production due to limitations on costs, labor, and time. To facilitate genetic change in livestock, traits that are typically not recorded, sex-limited, or only measurable in adulthood, but that are critical for important selection decisions early in life, are ideal candidates to consider for implementing whole-genome selection [21]. A significant number of genomic loci associated with age at menarche have been identified in humans through GWAS [22–24]. With the development of a commercially available, high-density SNP array for pig [25], genome-wide selection and the identification of genes and pathways affecting age at puberty of swine is feasible. Genome-wide associations for age at puberty [26] and delayed puberty [27] in the pig have previously been described. Several of the candidate genes associated with age at puberty are expressed in tissues from the hypothalamic-pituitary-ovarian axis [26] and are involved in sexual and social behavior, energy balance and oocyte maturation. Candidate genes associated with delayed puberty are similarly expressed but are more involved with synchronization of reproductive behavior and ovulation [27]. The objectives of this study were to use high-density genotyping and genome-wide association analysis to identify chromosomal regions and genes influencing age at puberty in a line of white-composite pigs and determine the proportion of genetic variance explained by the markers.

Methods

Care and handling of all animals included in this study was according to procedures outlined in *Guide for Care and Use of Agricultural Animals in Agricultural Research and Teaching* [28] and approved by the USMARC Animal Care and Use Committee.

Animals and data

A composite population was developed in 2001 using maternal and terminal Landrace, Duroc, and Yorkshire lines. Full- and half-sib matings were prevented, otherwise matings were random. Twelve original sire-lines were maintained and semen from all sire-lines was used to produce approximately 600 litters per generation. Additional details of the development of this population were previously reported [29]. Gilts born during 2005 through 2008 ($n = 759$) were genotyped and used for this study.

Estrus detection was performed daily from 140–240 days using 5–6 mature boars (>11 mo of age) placed in an alleyway between two pens of gilts, during which time herdsmen applied back pressure to gilts within each pen and observed them for estrous behavior. Age at puberty was defined as the date in which the first standing estrus was detected. Gilts that did not show signs of estrus by 240 days were harvested at the USMARC abattoir at an average age of 241 days, and the ovaries were inspected to determine whether they had not cycled and were classified as nonpubertal, or had cycled and were classified as behaviorally anestrus. Gilts that were observed to reach puberty before slaughter age had an average age of 200 ± 14.5 (mean ± SD) days at first estrus.

DNA isolation, SNP array genotyping and quality control

Genomic DNA was extracted from frozen tail tissue using the Wizard SV Genomic DNA Purification kit (Promega, Madison, WI, USA) for all phenotyped pigs. Samples of 300 ng at a concentration ≥ 75 ng/µl of DNA were genotyped using the Illumina PorcineSNP60 Bead-Chip containing 64,232 SNP (Illumina, San Diego, CA,USA) [25]. Genotypic reactions were completed at U.S. Meat Animal Research Center (USMARC, Clay Center, NE, USA) and then scanned at the USDA, ARS, Bovine Functional Genomics Laboratory (Beltsville, MD, USA). Scan results were interpreted at USMARC using Illumina's BeadStudio Genotyping software. Genotypes were called for 59,895 SNP spanning the entire porcine genome. Chromosome and position locations for each marker were according to the *Sus scrofa* genome assembly 10.2 (http://www.ncbi.nlm.nih.gov/projects/mapview/map_search.cgi?taxid=9823).

Any SNP with unknown chromosome positions, those located on SSCY, those with call rates < 95 %, or minor allele frequencies < 0.05 were excluded from the data set. Animals were eliminated ($n = 7$) if > 5 % of SNP were missing or for failing a Mendelian segregation (parentage) test. After utilizing these quality control measures, a total of 41,148 SNP out of a total of 64,232 SNP on the array qualified for GWAS. A genotypic principal components analysis was done using genotypes of all 759 phenotyped animals with SNP & Variation Suite v8.4.1 (Golden Helix, Inc., Bozeman, MT, http://goldenhelix.com/) [30]. The first three principal components are plotted in Additional file 1: Figure S1 and does not show population stratification due to breed.

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 3 of 9

Genome-wide association analyses

The analyses were implemented with a Bayes C model averaging approach using version 4.61 GENSEL software (http://archive.is/bigs.ansci.iastate.edu). Bayes C uses a common SNP variance that is reliably estimated from the data. The Bayes C has been explained previously by Kizilkaya et al. [31].

The following modified statistical model from Kizilkaya et al. [31] was used:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u} + \boldsymbol{e}$$

where $\mathbf{y}$ is the vector of phenotypes (age at puberty), $\mathbf{X}$ is an incidence matrix of fixed effects ($\boldsymbol{\beta}$), $\mathbf{Z}$ is a matrix of SNP genotypes that were fitted as random effects ($\boldsymbol{u}$) distributed N ($\mathbf{0}$, σ_u^2), and $\boldsymbol{e}$ is the vector of random residual effects assumed to be normally distributed N ($\mathbf{0}$, σ_e^2). A fixed classification factor used in this statistical model was year-season of birth as all females born in the same season were uniformly managed and monitored for first estrus.

Priors for variance components were taken from residual and additive genetic variance components from a preliminary analysis using MTDFREML [32]. The model fitted was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e}$$

where $\mathbf{y}$ represented a vector of observations; $\mathbf{b}$ was a vector of fixed effects; $\mathbf{a}$ was a vector of random additive genetic effects of animals, which was assumed to be distributed N ($\mathbf{0}$, $\mathbf{A}\sigma_a^2$), where $\mathbf{A}$ was the numerator relationship matrix among animals and $\mathbf{e}$ was a vector of residual effects, which was assumed to be distributed N ($\mathbf{0}$, $\mathbf{I}\sigma_e^2$) and where $\mathbf{I}$ was the identity matrix. Incidence matrix $\mathbf{X}$ related records to fixed effects and incidence matrix $\mathbf{Z}$ related records to additive genetic random effects. The fixed effects were those used in the Bayesian model above. The four generation pedigree file included 37 sires, 380 dams and 759 females with phenotypic data.

Bayes Cπ option of GENSEL was used to estimate $\boldsymbol{\pi}$ at 0.9935 where $\boldsymbol{\pi}$ is the prior probability that any SNP would have a zero effect. Bayes C was utilized for the analysis of SNP effects with a burn-in of 1000 iterations and a total of 51,000 iterations in a Markov chain. The results from this analysis included posterior distributions for the effects of each of the 41,148 markers, adjusted for the portfolio of all the other fitted marker effects in the model.

After the Bayes C analysis, the Predict option of GEN-SEL was used to estimate the genetic variance of sliding windows of five consecutive SNP assigned by chromosome-position order beginning with the first five SNP on SSC1 and ending with the last five SNP on SSCX. There were 8111 non-overlapping five-SNP windows available in the whole pig genome for statistical testing. Based on our estimate of π, 267 SNP effects were expected to be important (41,148 x (1-π) = 267). Therefore we submitted 267 5-SNP windows explaining the greatest proportion of genetic variance, defined as putative QTL, for statistical testing utilizing the Bootstrap option of GENSEL.

Bootstrap analysis for hypothesis testing

To construct the distribution of the test statistic (genetic variance of a five-SNP window) for each putative QTL, bootstrap samples were produced using the posterior means of the 41,148 SNP. This involved creating 1000 bootstrap data sets [33–35]. These bootstrap samples were constructed according to the null hypothesis of no QTL in the identified SNP window. Construction of the bootstrap begins with the results of the Bayes C analysis including the SNP effects of all markers except the SNP that are within the region of the putative QTL, which are set to zero, and the estimates of the fixed effects relevant to each animal's phenotypic record. Residual effects are sampled according to the residual variance previously found and added to each record. The only difference between bootstrap data replicates was due to different residual effects being sampled in different replicates. Each bootstrap sample was reanalyzed using the same Bayes C model used for the real data, and the genetic variances of the SNP window corresponding to the QTL were accumulated for comparison with the test statistic represented by the genetic variance of the SNP window identified in the analysis of the real data. If just one bootstrap statistic from the 1000 simulated exceeded the test statistic from the real data, the comparison-wise P-value was determined to be $0.001 < P < 0.002$.

Multiple testing was taken into account considering the proportion of false positives [36]. This approach controls the proportion of false positive conclusions across all tests undertaken, rather than the probability of making one mistake over all tests, as would be the interpretation of an experiment-wise error correction. The proportion of false positives is calculated as a function of the average comparison-wise Type I error rate, the proportion of true null hypotheses tested among all hypotheses tested, and the power of the test.

Candidate gene search

Intervals for candidate gene searches were defined as the 5-SNP window plus the 300 kb flanking regions in QTL peaks. Annotated genes contained in this window were identified using Ensembl BioMart tool (http://ensembl.org/Sus_scrofa/Info/Index) and the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway) with the Sus scrofa Build 10.2 assembly.

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 4 of 9

Linkage disequilibrium (LD) analysis

Linkage disequilibrium (r^2) was estimated for the SNP using Haploview 4.0 software [37] (http://www.broad.mit.edu/mpg/haploview/index.php). Haplotype blocks were based on pairwise LD values.

Results and discussion

In the present study, a GWAS using the PorcineSNP60 BeadChip was performed by means of Bayes C model averaging with random SNP effects for age at puberty. The variances, heritability, and the proportion of total variance explained by the markers are shown in Table 1. Using MTDFREML, heritability was found to be moderate at 0.319, similar to heritabilities of 0.38–0.46 that have been reported before [19, 20]. Rothschild and Bidanel [38] summarized age at puberty heritability estimates and determined a mean of 0.33 with a range of 0.00 to 0.64. The proportion of total variance explained by the markers was estimated by GENSEL to be 0.11 and indicates that the SNP explain less genetic variation than the infinitesimal model. While both measures of heritability should have similar expectations, the reason for variation between the two methods in unknown. In our previous studies on litter traits, similar variability was seen between pedigree-based and marker-based estimates of heritability [29]. Additional file 2: Table S1 presents information about the 222 5-SNP window associations identified in this study including percent of genetic variance (GV) explained by each QTL, and bootstrap *P*-values. Twenty seven QTL regions were significantly ($P < 0.01$) associated with age at puberty after bootstrap analysis (Table 2). Fifteen of these were the highest ranked for genetic variance. Statistical testing identified one QTL with $P < 0.001$, nine QTL with $0.001 < P < 0.005$, and 17 QTL with $0.005 < P \leq 0.010$. The QTL were identified on all autosomes as well as SSCX and explained 29.4 % of the GV identified by SNP markers tested. Sixteen additional putative QTL with $P \leq 0.027$ located in the vicinity of a significant QTL are also listed in Additional file 2: Table S1.

Twenty-eight of the QTL detected were within 1 Mb of other significant QTL (ranging from two to six 5-SNP windows within 1 Mb of a significant QTL). These adjacent groups of QTL may actually be due to a single quantitative trait nucleotide or gene in the region. High LD between markers in adjacent 5-SNP windows was observed in six QTL regions on SSC2 from 97.3–99.7 Mb, SSC6 from 88–89 Mb, SSC8 from 71.8–73.8 Mb, SSC9 from 93–96 Mb, SSC10 from 35.6–37.7 Mb and on SSC16 from 26–27.3 Mb (Additional file 3: Figure S2). The region on SSC2 at 98 Mb was previously associated with delayed puberty in pigs [27].

Sixteen of the QTL were within 0.5 Mb of previously reported associations [27] from the same population where attainment of puberty was considered a categorical trait and a case–control experimental design was implemented. Animals from the previous study that had not reached puberty were not included in the current study as their age at puberty was unknown. Two associations were located in published QTL for age of puberty in Meishan x European pig resource populations on SSC1 at 292 Mb [39] and on SSC10 at 68.5-69.5 Mb [40, 41]. Eight regions corresponded to those previously reported by Tart et al. [26] in a genome-wide association study (Additional file 2: Table S1).

Candidate genes in QTL regions

The most significant QTL accounting for 9.7 % of the genetic variance was located on SSC12 at 15 Mb within the growth hormone (*GH1*) and chorionic somatomammotropin hormone (*CSH1*) gene cluster. Growth hormone and its receptor are necessary for the onset and normal course for attainment of puberty [42]. There are age related changes in somatotropin secretion in pigs, with serum concentrations of GH declining as gilts approach pubertal age [43–45]. A second QTL accounting for over 7.1 % of the genetic variance was located on SSC7 at 75 Mb near the *PRKD1* and *C14orf 23* loci. *PRKD1* is associated with body-mass index (BMI) in humans [46] and an SNP in this window (INRA0026430) near *C14orf23* was previously reported to be associated with delayed puberty in relatives of the same population of pigs [27].

For regions previously reported by Tart et al. [26] likely candidate genes are *CRTC1* (SSC2:59 Mb), *PAPPA* (SSC1:288 Mb), *CCT6A* (SSC3:17 Mb), *CELF4* (SSC6:115 Mb), *EDEM3* (SSC9:139 Mb), *SMYD2* (SSC9:142 Mb), and *XXYLT1* (SSC13:142 Mb). *CCT6A* expression was shown to be higher in sexually mature laying hen ovaries than in immature ovaries [47].

Six regions contained genes associated with age at menarche in humans, and identified by Tart et al. [26] as well as the current study to be associated with age at puberty in pigs. *CRTC1* (SSC2:58.9 Mb) was associated with age at menarche in humans by Elks et al. [22]. This

Table 1 Summary statistics for age at puberty in pigs

Method	Number of Animals	Marker Heritability[1]	Genetic Variance	Residual Variance	Total Variance
MTDFREML	752	0.319 ± 0.102[2]	63.0	134.8	197.8
GenSel Bayes C	752	0.111	20.9	167.0	187.9

[1]Marker heritability (h^2) is the proportion of total variance explained by the markers
[2]Only the standard error for heritability was provided by MTDFREML and no standard errors were provided by GenSel

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 5 of 9

Table 2 Most significant QTL for age at puberty in pigs

SSC	Start[1]	End[1]	Marker 1	Marker 5	GenVar (%)	Rank[2]	p-value	-log(p)	Gene	Human trait or function
12	14,910,939	15,006,310	M1GA0016274	ASGA0053410	9.7332	1	0.0001	4.000	*GH1*	growth/height
6	68,823,187	68,890,361	ASGA0091283	ALGA0035583	0.3064	35	0.001	3.000	*TMEM51*	
1	291,202,322	291,277,150	ALGA0009830	M1GA0001507	0.6121	10	0.003	2.523	*BRINP1*	
3	26,517,667	26,631,496	MARC0007734	ALGA0018104	0.5091	15	0.003	2.523	*GPR139*	
7	43,190,587	43,400,560	ALGA0040857	ASGA0033098	0.1673	75	0.003	2.523	*UBR2*	obesity
7	75,121,635	75,460,294	INRA0026429	H3GA0022045	7.1362	2	0.003	2.523	*C14orf23, PRKD1*	BMI/obesity
15	42,527,231	43,052,985	H3GA0044224	ALGA0114567	1.7467	3	0.003	2.523	*ANGPT2*	ovulation
3	11,143,547	11,286,978	H3GA0008684	ASGA0013430	0.5400	13	0.004	2.398	*GTF2IRD1*	
9	142,232,369	142,338,853	ASGA0044888	ASGA0044901	1.4283	4	0.004	2.398	*SMYD2*	
10	18,347,386	18,665,600	ASGA0046792	DRGA0010326	0.3992	24	0.004	2.398	*SDCCAG8*	obesity
3	11,960,567	12,093,180	ALGA0017578	ALGA0017611	1.0870	6	0.005	2.301	*GATSL2*	
14	138,428,984	138,850,067	DRGA0014684	ALGA0082115	0.4572	19	0.005	2.301	*SLC18A2*	dopamine transport
1	121,626,075	121,978,687	M1GA0001099	ASGA0004239	0.4196	21	0.006	2.222	*RORA*	menarche
3	21,596,232	22,794,763	ASGA0013855	ASGA0094123	0.1765	70	0.006	2.222	*AQP8*	menarche
3	27,054,931	27,208,960	MARC0085816	ALGA0124353	1.1368	5	0.006	2.222	*GPRC5B*	BMI/obesity
6	127,194,648	127,655,907	MARC0001714	ASGA0029572	0.3900	25	0.006	2.222	*ACADM*	metabolite levels
15	148,542,045	148,722,557	ASGA0071543	ASGA0071569	0.1564	83	0.007	2.155	*HJURP*	
8	15,600,021	15,671,195	DRGA0008334	ASGA0037920	0.2543	41	0.008	2.097	*SLIT2*	
8	109,067,052	109,213,652	H3GA0025237	MARC0017963	0.2658	39	0.008	2.097	*TRPC3*	
11	55,668,610	55,844,999	ALGA0062350	ALGA0062355	0.3136	34	0.008	2.097	*RNF219*	
16	44,362,264	44,481,215	ALGA0090494	MARC0073104	0.2123	50	0.008	2.097	*IPO11*	
18	50,837,228	51,030,054	DRGA0017050	ASGA0105879	0.5142	14	0.008	2.097	*SNX10*	visceral adipose tissue
2	133,313,068	133,376,955	INRA0009763	MARC0038747	0.2935	36	0.009	2.046	*ZNF608*	BMI/obesity
10	40,710,979	40,868,229	MARC0038064	ALGA0058443	0.0985	149	0.009	2.046	*LINGO2*	BMI/obesity
15	142,893,866	142,992,475	MARC0112236	ALGA0087841	0.3401	29	0.009	2.046	*KPNA4*	
2	157,110,989	157,223,320	MARC0039166	M1GA0003373	0.2571	40	0.01	2.000	*ABLIM3*	
2	157,270,386	157,362,402	H3GA0008275	ALGA0017005	0.4738	17	0.01	2.000	*AFAP1L1*	obesity

[1]Start and End refer to SNP position in Sus scrofa Build 10.2
[2]Windows are ranked is by % genetic variance

gene also leads to infertility (failure to ovulate) in mice that lack a functional *CRTC1* [48] through mediation of the leptin-kisspeptin-GnRH pathway. Kisspeptin is a key regulator of LH secretion in the gilt [49]. *IQCH* located on SSC 1 at 183 Mb was associated with age at menarche in humans [22, 23]. In addition, three SNP in *IQCH* form a haplotype associated with a major gene affecting ovulation rate in twinner cattle [50]. *RORA* at 122 Mb on SSC1 was associated with age at menarche [22, 23] while a deletion in *RORA* in staggerer mice causes delayed puberty, reduced oocyte number and accelerated reproductive aging [51]. Finally, *AQP8* at 22 Mb on SSC3 [22] and *GPRC5B* at 27 Mb on SSC3 [24] have been associated with age at menarche while *NEGR1* at 130 Mb on SSC6 was associated with age at menarche and BMI [22, 46].

Several loci that are associated with anthropometric traits in humans such as body-mass index (BMI) or its components, body mass (obesity) and height were also associated with pubertal age in pigs (Additional file 2: Table S1). These include *MTCH2* and *ZNF608* on SSC2 at 16 Mb (BMI and obesity) [46, 52], and 133 Mb (BMI) [46, 52], *NEGR1* on SSC6 at 130 MB (obesity) [53], *ADAMTSL3* and *PRKD1* on SSC7 at 57 Mb (BMI and height) [54, 55] and 75 Mb [46], and *NPFFR2* on SSC8 at 72 Mb (BMI) [56]. Several human studies have shown relationships between the onset of puberty and adolescent growth rate [57], BMI [58] and obesity [59]. Using Mendelian randomization, Mumby et al. [60] showed a causative effect of increased BMI on earlier age at menarche in humans. Some studies in pigs have shown that gilts with higher growth rates [61, 62] or fed high energy diets [63] reach puberty earlier than lower growth rate or food restricted gilts [16]. Although faster growing gilts typically reach puberty earlier, there is little experimental evidence for a discrete level of body fatness

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 6 of 9

necessary for puberty to occur in commercial pigs when lean tissue growth is not limiting [61, 64]. Metabolic state at critical periods of development [65,66] or the degree of positive energy balance as determined by lifetime growth rate [62, 67] are considered more important determinants of pubertal age in gilts. Many loci associated with human age at menarche, BMI or obesity are involved in energy balance, such as *CRTC1* [48], *NEGR1* [53], *MTCH2* [68], *RORA* [69] and *TBC1D1* [70]. Variation in *TBC1D1* and *ZNF608* has been shown to be associated with fatness traits in pigs [71, 72]. An SNP window about 200 kb proximal to leptin (LEP) on SSC18 at 20.9 Mb was associated with age at puberty. Leptin levels and age at puberty in the pig are genetically correlated [20] and an increase in leptin seems to be permissive for pubertal development in the pig [73]. When lines of pigs were experimentally selected for increased lean percentage and less backfat thickness, gilts were older at puberty [74] owing to a general negative effect on the intensity of expressed estrous behavior [75, 76]. Reduced intensity of expressed estrous behavior would make it harder to detect pubertal estrus. Given that expression of estrus behavior was used to phenotype age at puberty in the current study, identification of loci linking age at puberty with body fatness or metabolism may not be directly due to activation of the hypothalamic-pituitary-gonadal axis per se, but rather a result of associated effects on sexual behavior.

Several other candidate genes were identified that are involved in pubertal development in other species. Two locations coincided with genes associated with puberty in cattle; *IGF1R* on SSC1 at 153 Mb [77] and *ESRRG* on SSC10 at 9 Mb [78]. In rats, hypothalamic blockade of *NELL2* (on SSC5 at 78.5 Mb) expression reduces GnRH release and results in delayed puberty [79, 80]. Alpha-fetoprotein (AFP) on SSC8 at 73.7 Mb is essential for female fertility [81] as mutant female mice fail to ovulate. Markers in this region on SSC8 showed high linkage disequilibrium that extended for about 2 Mb (71.8–73.8 Mb) and includes the genes *NPFFR2*, *ADAMTS3*, *ALB* and *AFP*. Substance P encoded by the tachykinin gene (*TAC1*) on SSC9 at 85 Mb and activin receptor-like kinase 7 (*ACVR1C*) on SSC15 at 71 Mb are associated with delayed puberty in female knockout mice [82, 83]. Substance P in found in the preoptic and suprachiasmatic areas in the medial forebrain of the pig [84], which are areas of the porcine hypothalamus that contain GnRH neurons [85]. Substance P is also found in the adenohypophysis of the pig [86] where it acts directly on porcine gonadotrope cells to modulate secretion of LH [87]. Moreover, substance P reportedly attenuated growth-hormone releasing hormone stimulated secretion of GH in primary cultures of porcine pituitary cells [88]. Given that puberty in the gilt is accompanied with a reduction in GH secretion and increased LH secretion [89], these data are strong supportive evidence for the association of substance P with age at puberty in the pig found in the current study.

Conclusions

Although rare monogenic mutations can disrupt normal pubertal development in humans through the GnRH axis, most loci contain common genetic variants that contribute to variation in pubertal timing and are involved in co-regulation of height, BMI or obesity (reviewed in Day et al.) [90]. Results of the current study illustrate the complex relationship of growth and metabolism with puberty in the pig. Several of the loci identified to be associated with age at puberty in gilts in this study contain candidate genes that can have direct actions at the anterior pituitary gland or interact centrally with the kisspeptin-GnRH neural network to control secretion of LH and many loci were identified that have a physiological role for the onset of puberty in rodents, cattle and pigs and a genetic basis for sexual maturation in humans. This provides strong correlative evidence that these genomic associations in the pig are reliable. The genomic markers identified in the current experiment are an important resource that will be used to develop validated markers for age at puberty in commercial populations of pigs. Because age at puberty is not routinely measured or selected for in commercial herds, and is a predictive factor for sow longevity [41, 91] and lifetime productivity [16], these genomic markers will facilitate the use of genomic or marker-assisted selection to improve these traits.

Additional files

Additional file 1: Figure S1. Genotypic principal components analysis of 759 phenotyped animals; the first three principal components are plotted. (PDF 197 kb)

Additional file 2: Table S1. Genomic regions, SNP results and candidate genes associated with age at puberty in pigs. (XLSX 37 kb)

Additional file 3: Figure S2. Linkage disequilibrium plots of QTL regions of more than one consecutive 5-SNP window greater that 1 Mb with high r^2 values. (PDF 821 kb)

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 7 of 9

References

1. Barb CR, Hausman GJ, Rekaya R. Gene expression in the brain-pituitary adipose tissue axis and luteinising hormone secretion during pubertal development in the gilt. Soc Reprod Fertil Suppl. 2006;62:33–44.
2. Choi J-H, Yoo H-W. Control of puberty: genetics, endocrinology, and environment. Curr Opin Endocrinol Diabetes Obes. 2013;20:62–8.
3. Ojeda SR, Lomniczi A. Puberty in 2013: Unravelling the mystery of puberty. Nat Rev Endocrinol. 2014;10:67–9.
4. Peeters PH, Verbeek AL, Krol A, Matthyssen MM, de Waard F. Age at menarche and breast cancer risk in nulliparous women. Breast Cancer Res Treat. 1995;33:55–61.
5. Hunter DJ, Spiegelman D, Adami H-O, van den Brandt PA, Folsom AR, Goldbohm RA, et al. Non-dietary factors as risk factors for breast cancer, and as effect modifiers of the association of fat intake and risk of breast cancer. Cancer Causes Control. 1997;8:49–56.
6. Kotsopoulos J, Lubinski J, Lynch HT, Neuhausen SL, Ghadirian P, Isaacs C, et al. Age at menarche and the risk of breast cancer in *BRCA1* and *BRCA2* mutation carriers. Cancer Causes Control. 2005;16:667–74.
7. Brinton LA, Berman ML, Mortel R, Twiggs LB, Barrett RJ, Wilbanks GD, et al. Reproductive, menstrual, and medical risk factors for endometrial cancer: results from a case-control study. Am J Obstet Gynecol. 1992;167:1317–25.
8. Purdie DM, Green AC. Epidemiology of endometrial cancer. Best Pract Res Clin Obstet Gynaecol. 2001;15:341–54.
9. Sharma K, Talwar I, Sharma N. Age at menarche in relation to adult body size and physique. Ann Hum Biol. 1988;15:431–4.
10. Biro FM, McMahon RP, Striegel-Moore R, Crawford PB, Obarzanek E, Morrison JA, et al. Impact of timing of pubertal maturation on growth in black and white female adolescents: The National Heart, Lung, and Blood Institute Growth and Health Study. J Pediatr. 2001;138:636–43.
11. Okasha M, McCarron P, McEwen J, Smith GD. Age at menarche: secular trends and association with adult anthropometric measures. Ann Hum Biol. 2001;28:68–78.
12. Presser HB. Age at menarche, socio-sexual behavior, and fertility. Soc Biol. 1978;25:94–101.
13. Komura H, Miyake A, Chen CF, Tanizawa O, Yoshikawa H. Relationship of age at menarche and subsequent fertility. Eur J Obstet Gynecol Reprod Biol. 1992;44:201–3.
14. Patterson JL, Beltranena E, Foxcroft GR. The effect of gilt age at first estrus and breeding on third estrus on sow body weight changes and long-term reproductive performance. J Anim Sci. 2010;88:2500–13.
15. Saito H, Sasaki Y, Koketsu Y. Associations between age of gilts at first mating and lifetime performance or culling risk in commercial herds. J Vet Med Sci. 2011;73:555–9.
16. Miller PS, Moreno R, Johnson RK. Effects of restricting energy during the gilt developmental period on growth and reproduction of lines differing in lean growth rate: Responses in feed intake, growth, and age at puberty. J Anim Sci. 2011;89:342–54.
17. Young MG, Tokach MD, Aherne FX, Dritz SS, Goodband RD, Nelssen JL, et al. Effect of space allowance during rearing and selection criteria on performance of gilts over three parities in a commercial swine production system. J Anim Sci. 2008;86:3181–93.
18. Knauer MT, Cassady JP, Newcom DW, See MT. Phenotypic and genetic correlations between gilt estrus, puberty, growth, composition, and structural conformation traits with first-litter reproductive measures. J Anim Sci. 2011;89:935–42.
19. Rosendo A, Druet T, Gogué J, Canario L, Bidanel JP. Correlated responses for litter traits to six generations of selection for ovulation rate or prenatal survival in French Large White pigs. J Anim Sci. 2007;85:1615–24.
20. Kuehn LA, Nonneman DJ, Klindt JM, Wise TH. Genetic relationships of body composition, serum leptin, and age at puberty in gilts. J Anim Sci. 2009;87:477–83.
21. Goddard ME, Hayes BJ, Meuwissen THE. Genomic selection in livestock populations. Genet Res. 2010;92:413–21.
22. Elks CE, Perry JRB, Sulem P, Chasman DI, Franceschini N, He C, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nat Genet. 2010;42:1077–85.
23. Demerath EW, Liu C-T, Franceschini N, Chen G, Palmer JR, Smith EN, et al. Genome-wide association study of age at menarche in African-American women. Hum Mol Genet. 2013;22:3329–46.
24. Fernández-Rhodes L, Demerath EW, Cousminer DL, Tao R, Dreyfus JG, Esko T, et al. Association of adiposity genetic variants with menarche timing in 92,105 women of European descent. Am J Epidemiol. 2013;178:451–60.
25. Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, et al. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS ONE. 2009;4:e6524.
26. Tart JK, Johnson RK, Bundy JW, Ferdinand NN, McKnite AM, Wood JR, et al. Genome-wide prediction of age at puberty and reproductive longevity in sows. Anim Genet. 2013;44:387–97.
27. Nonneman D, Lents C, Rohrer G, Rempel L, Vallet J. Genome-wide association with delayed puberty in swine. Anim Genet. 2014;45:130–2.
28. FASS. (Federation of Animal Science Societies). *Guide for the Care and Use of Agricultural Animals in Research and Teaching, Third Edition*. Champaign, IL: Federation of Animal Science Societies; 2010.
29. Schneider JF, Rempel LA, Rohrer GA. Genome-wide association study of swine farrowing traits. Part I: Genetic and genomic parameter estimates. J Anim Sci. 2012;90:3353–9.
30. SNP & Variation Suite™ (Version 8.4) [Genotypic Principal Components Analysis]. Bozeman, MT: Golden Helix, Inc. Available from http://goldenhelix.com.
31. Kizilkaya K, Fernando RL, Garrick DJ. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. J Anim Sci. 2010;88:544–51.
32. Boldman KG, Kriese LA, Van Vleck LD, Van Tassell CP, Kachman SD. A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances. USDA, ARS, Lincoln, Nebraska, USA. 1995;(available: http://aipl.arsusda.gov/software/mtdfreml/ accessed 19 Aug 2015).
33. Fan B, Onteru SK, Du Z-Q, Garrick DJ, Stalder KJ, Rothschild MF. Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. PLoS ONE. 2011;6:e14726.
34. Onteru SK, Fan B, Nikkilä MT, Garrick DJ, Stalder KJ, Rothschild MF. Whole-genome association analyses for lifetime reproductive traits in the pig. J Anim Sci. 2011;89:988–95.
35. Onteru SK, Fan B, Du Z-Q, Garrick DJ, Stalder KJ, Rothschild MF. A whole-genome association study for pig reproductive traits. Anim Genet. 2012;43:18–26.
36. Fernando RL, Nettleton D, Southey BR, Dekkers JCM, Rothschild MF, Soller M. Controlling the proportion of false positives in multiple dependent tests. Genetics. 2004;166:611–9.
37. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21:263–5.
38. Rothschild MF, Bidanel J-P. Biology and Genetics of Reproduction. In: Rothschild MF, Ruvinsky A, editors. The Genetics of the Pig. Cambridge, UK: CABI, University Press; 1998. p. 313–43.

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 8 of 9

39. Bidanel JP, Rosendo A, Iannuccelli N, Riquet J, Gilbert H, Caritez JC, et al. Detection of quantitative trait loci for teat number and female reproductive traits in Meishan × Large White F2 pigs. Animal. 2008;2:813–20.

40. Rohrer GA, Ford JJ, Wise TH, Vallet JL, Christenson RK. Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. J Anim Sci. 1999;77:1385–91.

41. Nonneman DJ, Wise TH, Ford JJ, Kuehn LA, Rohrer GA. Characterization of the aldo-keto reductase 1C gene cluster on pig chromosome 10: possible associations with reproductive traits. BMC Vet Res. 2006;2:28.

42. Spiliotis BE. Growth hormone insufficiency and its impact on ovarian function. Ann NY Acad Sci. 2003;997:77–84.

43. Machlin LJ, Horino M, Hertelendy F, Kipnis DM. Plasma growth hormone and insulin levels in the pig. Endocrinology. 1968;82:369–76.

44. Klindt J, Stone RT. Porcine growth hormone and prolactin: concentrations in the fetus and secretory patterns in the growing pig. Growth. 1984;48:1–15.

45. Dubreuil P, Pelletier G, Petitclerc D, Lapierre H, Couture Y, Brazeau P, et al. Influence of age and sex on basal secretion of growth hormone (GH) and on GH-induced release by porcine GH-releasing factor pGRF(1-29NH$_2$) in growing pigs. Domest Anim Endocrinol. 1987;4:299–307.

46. Graff M, Ngwa JS, Workalemahu T, Homuth G, Schipf S, Teumer A, et al. Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. Hum Mol Genet. 2013;22:3597–607.

47. Kang L, Zhang Y, Zhang N, Zang L, Wang M, Cui X, et al. Identification of differentially expressed genes in ovaries of chicken attaining sexual maturity at different ages. Mol Biol Rep. 2012;39:3037–45.

48. Altarejos JY, Goebel N, Conkright MD, Inoue H, Xie J, Arias CM, et al. The Creb1 coactivator Crtc1 is required for energy balance and fertility. Nat Med. 2008;14:1112–7.

49. Lents CA, Heidorn NL, Barb CR, Ford JJ. Central and peripheral administration of kisspeptin activates gonadotropin but not somatotropin secretion in prepubertal gilts. Reproduction. 2008;135:879–87.

50. Kirkpatrick BW, Morris CA. A major gene for bovine ovulation rate. PLoS ONE. 2015;10:e0129025.

51. Guastavino JM, Larsson K. The staggerer gene curtails the reproductive life span of females. Behav Genet. 1992;22:101–12.

52. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42:937–48.

53. Boender AJ, van Gestel MA, Garner KM, Luijendijk MCM, Adan RAH. The obesity-associated gene *Negr1* regulates aspects of energy balance in rat hypothalamic areas. Physiol Rep. 2014;2:e12083.

54. Liu JZ, Medland SE, Wright MJ, Heath AC, Madden PAF, Duncan A, et al. Genome-wide association study of height and body mass index in Australian twin families. Twin Res Hum Genet. 2010;13:179–93.

55. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. 2008;40:575–83.

56. Hunt SC, Hasstedt SJ, Xin Y, Dalley BK, Milash BA, Yakobson E, et al. Polymorphisms in the *NPY2R* gene show significant associations with BMI that are additive to *FTO*, *MC4R*, and *NPFFR2* gene effects. Obesity. 2011;19:2241–7.

57. Cousminer DL, Berry DJ, Timpson NJ, Ang W, Thiering E, Byrne EM, et al. Genome-wide association and longitudinal analyses reveal genetic loci linking pubertal height growth, pubertal timing and childhood adiposity. Hum Mol Genet. 2013;22:2735–47.

58. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. Relation of age at menarche to race, time period, and anthropometric dimensions: The Bogalusa Heart Study. Pediatrics. 2002;110:e43.

59. Wagner IV, Sabin MA, Pfäffle RW, Hiemisch A, Sergeyev E, Körner A, et al. Effects of obesity on human sexual development. Nat Rev Endocrinol. 2012;8:246–54.

60. Mumby HS, Elks CE, Li S, Sharp SJ, Khaw K-T, Luben RN, et al. Mendelian randomisation study of childhood BMI and early menarche. J Obes. 2011;2011:180729.

61. Bortolozzo FP, Bernardi ML, Kummer R, Wentz I. Growth, body state and breeding performance in gilts and primiparous sows. Soc Reprod Fertil Suppl. 2009;66:281–91.

62. Kummer R, Bernardi ML, Schenkel AC, Amaral Filha WS, Wentz I, Bortolozzo FP. Reproductive performance of gilts with similar age but with different

63. growth rate at the onset of puberty stimulation. Reprod Domest Anim. 2009;44:255–9.

63. Zhuo Y, Zhou D, Che L, Fang Z, Lin Y, Wu D. Feeding prepubescent gilts a high-fat diet induces molecular changes in the hypothalamus-pituitary-gonadal axis and predicts early timing of puberty. Nutrition. 2014;30:890–6.

64. Patterson JL, Ball RO, Willis HJ, Aherne FX, Foxcroft GR. The effect of lean growth rate on puberty attainment in gilts. J Anim Sci. 2002;80:1299–310.

65. Rozeboom DW, Pettigrew JE, Moser RL, Cornelius SG, El Kandelgy SM. Body composition of gilts at puberty. J Anim Sci. 1995;73:2524–31.

66. Lents CA, Rempel LA, Klindt J, Wise T, Nonneman D, Freking BA. The relationship of plasma urea nitrogen with growth traits and age at first estrus in gilts. J Anim Sci. 2013;91:3137–42.

67. Beltranena E, Aherne FX, Foxcroft GR, Kirkwood RN. Effects of pre- and postpartum feeding on production traits at first and second estrus in gilts. J Anim Sci. 1991;69:886–93.

68. Bauer F, Elbers CC, Adan RAH, Loos RJF, Onland-Moret NC, Grobbee DE, et al. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. Am J Clin Nutr. 2009;90:951–9.

69. Crumbley C, Wang Y, Banerjee S, Burris TP. Regulation of expression of citrate synthase by the retinoic acid receptor-related orphan receptor α (RORα). PLoS ONE. 2012;7:e33804.

70. Szekeres F, Chadt A, Tom RZ, Deshmukh AS, Chibalin AV, Björnholm M, et al. The Rab-GTPase-activating protein TBC1D1 regulates skeletal muscle glucose metabolism. Am J Physiol Endocrinol Metab. 2012;303:E524–33.

71. Fontanesi L, Colombo M, Tognazzi L, Scotti E, Buttazzoni L, Dall'Olio S, et al. The porcine *TBC1D1* gene: mapping, SNP identification, and association study with meat, carcass and production traits in Italian heavy pigs. Mol Biol Rep. 2011;38:1425–31.

72. Rothammer S, Kremer PV, Bernau M, Fernandez-Figares I, Pfister-Schär J, Medugorac I, et al. Genome-wide QTL mapping of nine body composition and bone mineral density traits in pigs. Genet Sel Evol. 2014;46:68.

73. Barb CR, Hausman GJ, Czaja K. Leptin: A metabolic signal affecting central regulation of reproduction in the pig. Domest Anim Endocrinol. 2005;29:186–92.

74. Rydhmer L, Eliasson-Selling L, Johansson K, Stern S, Andersson K. A genetic study of estrus symptoms at puberty and their relationship to growth and leanness in gilts. J Anim Sci. 1994;72:1964–70.

75. Eliasson L, Rydhmer L, Einarsson S, Andersson K. Relationships between puberty and production traits in the gilt. 1. Age at puberty. Anim Reprod Sci. 1991;25:143–54.

76. Cameron ND, Curran MK. Responses in carcass composition to divergent selection for components of efficient lean growth rate in pigs. Anim Sci. 1995;61:347–59.

77. Fortes MRS, Li Y, Collis E, Zhang Y, Hawken RJ. The IGF1 pathway genes and their association with age of puberty in cattle. Anim Genet. 2013;44:91–5.

78. Fortes MRS, Reverter A, Zhang Y, Collis E, Nagaraj SH, Jonsson NN, et al. Association weight matrix for the genetic dissection of puberty in beef cattle. Proc Natl Acad Sci USA. 2010;107:13642–7.

79. Ha CM, Choi J, Choi EJ, Costa ME, Lee BJ, Ojeda SR. NELL2, a neuron-specific EGF-like protein, is selectively expressed in glutamatergic neurons and contributes to the glutamatergic control of GnRH neurons at puberty. Neuroendocrinology. 2008;88:199–211.

80. Ryu BJ, Kim HR, Jeong JK, Lee BJ. Regulation of the female rat estrous cycle by a neural cell-specific epidermal growth factor-like repeat domain containing protein, NELL2. Mol Cells. 2011;32:203–7.

81. Gabant P, Forrester L, Nichols J, Van Reeth T, De Mees C, Pajack B, et al. Alpha-fetoprotein, the major fetal serum protein, is not essential for embryonic development but is required for female fertility. Proc Natl Acad Sci USA. 2002;99:12865–70.

82. Simavli S, Thompson IR, Maguire CA, Gill JC, Carroll RS, Wolfe A, et al. Substance P regulates puberty onset and fertility in the female mouse. Endocrinology. 2015;156:2313–22.

83. Sandoval-Guzmán T, Göngrich C, Moliner A, Guo T, Wu H, Broberger C, et al. Neuroendocrine control of female reproductive function by the activin receptor ALK7. FASEB J. 2012;26:4966–76.

84. Kumar MSA, Becker T, Ebert K. Distribution of substance P, GnRH, Met- enkephalin in the central nervous system of the pig. Brain Res Bull. 1991;26:511–4.

85. Kineman RD, Leshin LS, Crim JW, Rampacek GB, Kraeling RR. Localization of luteinizing hormone-releasing hormone in the forebrain of the pig. Biol Reprod. 1988;39:665–72.

Nonneman *et al. BMC Genetics* (2016) 17:50

Page 9 of 9

86. Yanaihara C, Sato H, Hirohashi M, Sakagami M, Yamamoto K, Hashimoto T, et al. Substance P radioimmunoassay using N^a-tyrosyl-substance P and demonstration of the presence of substance P-like immunoreactivities in human blood and porcine tissue extracts. Endocrinol Jpn. 1976;23:457–63.

87. Hidalgo-Díaz C, Castaño JP, López-Pedrera R, Malagón MM, García-Navarro S, Gracia-Navarro F. A modulatory role for substance P on the regulation of luteinizing hormone secretion by cultured porcine gonadotrophs. Biol Reprod. 1998;58:678–85.

88. Glavaski-Joksimovic A, Jeftinija K, Jeremic A, Anderson LL, Jeftinija S. Mechanism of action of the growth hormone secretagogue, L-692,585, on isolated porcine somatotropes. J Endocrinol. 2002;175:625–36.

89. Barb CR, Hausman GJ, Lents CA. Energy metabolism and leptin: Effects on neuroendocrine regulation of reproduction in the gilt and sow. Reprod Domest Anim. 2008;43 Suppl 2:324–30.

90. Day FR, Perry JRB, Ong KK. Genetic regulation of puberty timing in humans. Neuroendocrinology. 2015. doi: 10.1159/000431023.

91. Knauer M, Stalder KJ, Serenius T, Baas TJ, Berger PJ, Karriker L, et al. Factors associated with sow stayability in 6 genotypes. J Anim Sci. 2010;88:3486–92.