BMC
Genomics

High-resolution association mapping of number of teats in pigs reveals regions controlling vertebral development

Naomi Duijvesteijn[1*], Jacqueline M Veltmaat[2], Egbert F Knol[1] and Barbara Harlizius[1]

Abstract

Background: Selection pressure on the number of teats has been applied to be able to provide enough teats for the increase in litter size in pigs. Although many QTL were reported, they cover large chromosomal regions and the functional mutations and their underlying biological mechanisms have not yet been identified. To gain a better insight in the genetic architecture of the trait number of teats, we performed a genome-wide association study by genotyping 936 Large White pigs using the Illumina PorcineSNP60 Beadchip. The analysis is based on deregressed breeding values to account for the dense family structure and a Bayesian approach for estimation of the SNP effects.

Results: The genome-wide association study resulted in 212 significant SNPs. In total, 39 QTL regions were defined including 170 SNPs on 13 *Sus scrofa* chromosomes (SSC) of which 5 regions on SSC7, 9, 10, 12 and 14 were highly significant. All significantly associated regions together explain 9.5% of the genetic variance where a QTL on SSC7 explains the most genetic variance (2.5%). For the five highly significant QTL regions, a search for candidate genes was performed. The most convincing candidate genes were *VRTN* and *Prox2* on SSC7, *MPP7*, *ARMC4*, and *MKX* on SSC10, and vertebrae *δ-EF1* on SSC12. All three QTL contain candidate genes which are known to be associated with vertebral development. In the new QTL regions on SSC9 and SSC14, no obvious candidate genes were identified.

Conclusions: Five major QTL were found at high resolution on SSC7, 9, 10, 12, and 14 of which the QTL on SSC9 and SSC14 are the first ones to be reported on these chromosomes. The significant SNPs found in this study could be used in selection to increase number of teats in pigs, so that the increasing number of live-born piglets can be nursed by the sow. This study points to common genetic mechanisms regulating number of vertebrae and number of teats.

Keywords: Pigs, Number of teats, Vertebrae, Genome-wide association study, Somites, De-regressed breeding values

Background

A favorable genetic trend for total number of piglets born has been observed in the last decade [1,2]. Therefore, selection pressure on the number of teats has been applied in order to provide enough teats for the larger litters. Given the heritable nature of variation in number of teats in sheep as described by Bell in 1898 [3], and cases of familial supernumerary breasts (polymastia) or nipples (polythelia) in humans [4], genomic *loci* affecting teat number must exist. Indeed, the use of Best Linear Unbiased Prediction (BLUP) for estimating breeding values (EBVs) using phenotypes of both females and

males, has resulted in an increase of the number of teats (data not shown) and heritability estimates are moderate with estimates between 0.2 and 0.47 [5,6]. Besides the use of quantitative genetics to select the best sows, many studies have used genetic markers - mainly microsatellites - to identify QTL (Quantitative Trait Loci). The QTL studies on number of teats (NTE) listed in the Pig QTL database [7,8] report QTL on all porcine chromosomes except SSC9, SSC13, SSC14, SSC18 and SSCY. Although many QTL were reported [9-18], they cover large regions and the functional mutations and underlying biological mechanisms have not yet been identified. Interestingly, some QTL for NTE seem to overlap with those for number of vertebrae [11].

Teats (or nipples) develop each as an appendage to previously formed mammary gland rudiments (MRs) during

* Correspondence: naomi.duijvesteijn@TOPIGS.com
[1]TOPIGS Research Center IPG, PO Box 43, 6640AA Beuningen, The Netherlands
Full list of author information is available at the end of the article

pre-natal life [19,20]. Therefore, the number of teats correlates with the number of MRs induced and maintained at least until teat formation. The number of mammary glands varies among mammalian species, but even in humans, who normally form one pair of breasts, there are at least 6 other positions that additional breasts can randomly occupy on either side of the body [21]. Their positions range from armpit (axilla) to groin (inguen), thus span the same region where pigs form their mammary glands and teats. On both sides lateral to the ventral midline, one can draw imaginary fluent lines from both axillae to both inguenae, called mammary lines or milk lines. During embryonic life, these lines exist as histologically and molecularly distinct bands in the surface ectoderm, connecting all positions where mammary glands may form on either side of the body in any given mammalian species [22-24].

Studies on genetically engineered mice (GEMs) have revealed some insights in genetic and cellular mechanisms of mammary gland and nipple formation. Whereas wild type mice normally form five pairs of mammary glands along the mammary lines [22], modification of certain genes can alter this number [25]. For example, loss of either *Nrg3*, *Pax3*, *Gli3*, *Fgf10*, or *Hoxc6* [26-29] abolishes the formation of different, gene-specific, subsets of MRs [25]. Deletion of *e.g. p63* [30,31] or *Tbx3* (T-box gene 3) [32] may abolish the formation of all five MRs, while reduction of Wnt signaling by means of deletion of *e.g. Lef1* (Lymphoid enhancer factor 1) [33] or *Pygo2* (Pygopus 2) allows MR induction but leads to MR regression prior to nipple formation [34]. Even if the MRs are maintained, nipple formation may not occur due to *e.g.* a lack of PTHrP (parathyroid hormone related peptide) signaling [35]. Conversely, ectodermal overexpression of the genes encoding the receptor ligands Eda-A1 (Ectodysplasin-A1) [36] or *Nrg3* [37], or suppression of the Wnt-antagonists *Lrp4* or Wise/ Sostdc1, may lead to formation of one or several supernumerary mammary glands in a gene-specific pattern or region along the mammary line [38,39]. These data provide evidence for genetic determinants for the number of mammary glands, and moreover, for the positional variation in activity of, or requirement for, these genetic determinants along the mammary line [25].

Studies carried out in the 1960's on rabbit and mouse embryos had revealed that mammary gland formation in the surface ectoderm is initiated by factors in the dermal mesenchyme underlying the surface ectoderm [27,40,41]. The dermal mesenchyme is derived from the somites, which are also the precursor structures for the vertebrae and ribs. Interestingly, some of the genes mentioned above (*e.g. Gli3*, *Fgf10*, *Nrg3*, *Pax3*, *Hoxc6*), are in wild type mice expressed in the somites and/or in the dermal mesenchyme. It is now known that induction of the third pair of MRs in mice, MR3, depends on Gli3-mediated *Fgf10* expression in the somites [27]. Somitic expression of Raldh2, an enzyme involved in retinoic acid synthesis, has also been associated with induction of MR3 [42]. With this relationship between somitic gene expression and the number of MRs induced in mouse embryos; and knowing that somites also give rise to the vertebrae, we may expect that some QTL and candidate genes for NTE in pigs are associated with the number of vertebrae in pigs, or vertebrae development in mammalian species in general.

The availability of the Illumina PorcineSNP60 Beadchip [43] allowed us to perform a Genome-Wide Association Study (GWAS) for NTE by genotyping 936 Large White pigs. The better coverage of the whole pig genome by high-density Single Nucleotide Polymorphisms (SNPs) on these Beadchips, combined with advanced statistical methods can improve fine mapping of specific QTL as demonstrated previously for other traits [44-46]. In the current study, we based our analysis on deregressed breeding values [47] to account for the dense family structure and a Bayesian approach for estimation of the SNP effects. A total of five highly significant QTL were identified at high resolution including new regions on Sus scrofa chromosome (SSC) 9 and SSC14. Interestingly, in three of these regions we identified genes regulating vertebrae development as candidate genes determining the number of teats.

Results

Identification of QTL for teat number

In total 949 pigs were originally genotyped, but 13 animals were removed during quality control of the data. The remaining 936 animals had an average call rate of 0.995. The mean teat number was 15.3 (SD = 0.94) with a minimum of 14 and a maximum of 19 teats. A minimum of 14 teats is used as a breeding decision, which results in a slightly skewed distribution of the phenotype in this dataset. The (deregressed) EBVs are normally distributed (results not shown). The distribution of the weighting factors calculated for NTE to account for heterogeneous variances are shown in Figure 1. The two distributions observed reflect the deviation between animals with or without offspring information. Animals with a weighting factor below 1 have no offspring where the animals above one have on average 143 offspring with records on NTE. The estimated heritability was 0.42. The GWAS resulted in 212 SNPs with a BF >10 of which 6 SNPs had a BF > 100 (Figure 2). In total 39 QTL regions were defined containing 170 SNPs on 13 chromosomes. One candidate QTL region showed an r^2 higher than 0.7 with another candidate QTL where the distance was over 2 Mb (4 Mb respectively). These regions were combined into one region (QTL number 30 on SSC14).
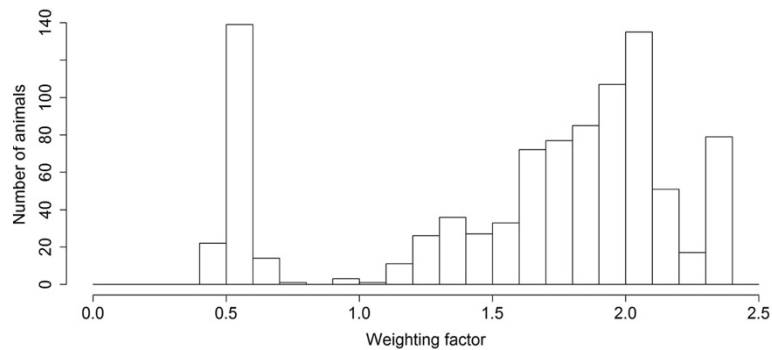
Figure 1 Distribution of the weights of the breeding values calculated according to the methodology by Garrick et al. [47]. The x-axis shows the weighting factors used to account for differences in phenotypic information density for each animal.

The 39 significantly associated regions explain 9.5% of the genetic variance with QTL number 13 on SSC7 explaining most of the genetic variance (2.5%, see Table 1). The most significant SNP within this region is ALGA0122954. The allele substitution effect of the A allele (major allele) to the G allele constitutes an additional 0.21 teat on a phenotypic level. This is an indication for overdominance because the heterozygous animals have on average 15.45 teats compared to 15.16 for AA animals and 15.37 for GG animals, respectively. The second largest QTL located on SSC12 explains more than 1% of the genetic variance, with ALGA0066876 being the SNP which explains most of the variance within this region. The allele substitution effect of the A allele (major allele) to the G allele is almost 0.09 teat indicating that the effect is completely additive (see Table 1). In total there are 14 QTL which each explain more than 0.2% of the genetic variance. In Figure 3, the explained variance per chromosome is shown. SSC7, which contains the largest QTL, is also the chromosome which explains most of the genetic variance. The second largest contributing chromosome is SSC1, but this is only due to the large number of

SNPs all contributing a small variance ($\sigma_{g_0}^2$) because SSC1 is by far the longest autosome in pigs. The total attributed SNP variance when placed in the null distribution was 69%.

Gene identification within QTL regions

Within the QTL regions, 489 genes (unique Ensembl Gene IDs) were mapped (see Additional file 1 for the genes). For the five highly significant QTL regions on SSC7, 9, 10, 12 and 14 with a BF near or larger than 100, a search for candidate genes was performed. On SSC7 between 102 and 105 Mb, a gene named *Vertnin (VRTN)* and encoding a potential DNA binding factor, is located at 103.45 Mb. The number-increase allele (Q) of this gene, was shown to add an additional thoracic segment to the pig compared to the wild-type (*wt*) [48]. Among the 14 genes annotated to the QTL area on SSC12 explaining more than 1% of the genetic variance, is *δ-EF1* mapping to 51.96 Mb. We consider this the most likely candidate gene, because it encodes transcriptional repressor involved in skeletal development [49]. On SSC10, three candidate genes were identified
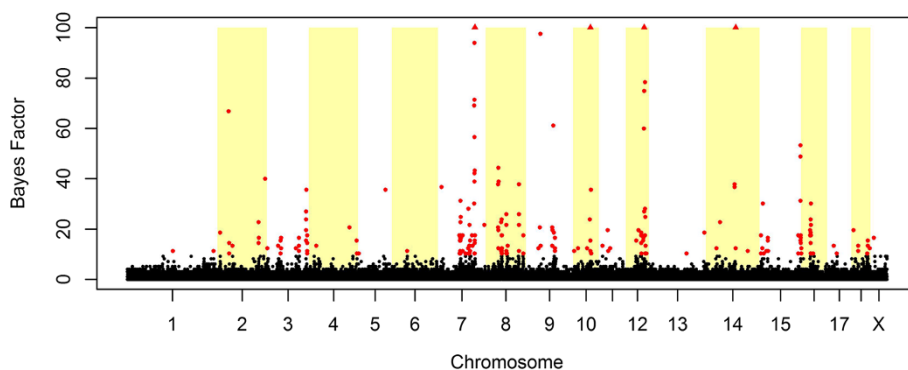

Figure 2 Genome-wide association for number of teats (NTE) in 936 purebred pigs from a Large White population. On the x-axis are the physical positions of the SNPs by chromosome. On the y-axis are Bayes Factors (BF) per SNP shown. BF <10 are black dots, BF >10 are red dots and BF > 100 are red triangles. The y-axis is cut-off at a BF of 100.

Table 1 Significant QTL regions per chromosome (SSC) associated with number of teats including the most significant SNPs, minor allele frequency (MAF), Bayes Factor (BF), allele substitution effect, the genetic variance explained by the region and the number of genes (Ensembl Gene IDs) found within the region

QTL number	SSC	Region[1]	Most sign. SNP	MAF	BF	Allele subst. effect[2]	% Gen. var. explained region[3]	Found genes
1	2	32.99-34.51	ALGA0012906	0.27	66.7	−0.143	0.17	1
2	2	140.44-140.93	ISU10000081	0.28	22.7	−0.17	0.11	12
3	3	39.13-39.28	MARC0000538	0.35	15.4	−0.079	0.04	1
4	3	44.55-45.82	ASGA0014296	0.29	16.4	−0.067	0.06	22
5	3	113.1-113.41	ALGA0020641	0.3	16.4	−0.068	0.08	1
6	3	134.97-135.69	ASGA0090006	0.23	35.5	−0.051	0.23	12
7	3	138.21-138.59	ALGA0115665	0.18	17.5	−0.117	0.07	0
8	4	139.72-140.10	H3GA0014860	0.39	15.4	−0.002	0.04	3
9	7	55.41-58.19	H3GA0021793	0.35	31.2	0.187	0.57	28
10	7	63.59-63.78	DRGA0007689	0.46	17.5	0.151	0.05	4
11	7	73.7-74.94	DRGA0007771	0.32	10.2	−0.155	0.04	1
12	7	87.05-91.15	ALGA0042950	0.35	28	0.157	0.29	7
13	7	102.01-105.22	ALGA0122954	0.21	210.1	0.227	2.51	56
14	8	33.26-35.00	ASGA0093882	0.34	44.3	0.16	0.38	11
15	8	43.92-45.19	ALGA0119079	0.23	22.7	−0.205	0.06	12
16	8	55.18-58.63	ALGA0047889	0.18	23.8	−0.198	0.24	26
17	8	80.00-81.42	MARC0044036	0.18	25.9	0.246	0.12	8
18	8	130.69-130.91	ALGA0049466	0.46	37.7	0.054	0.2	1
19	8	143.24-143.35	ALGA0050050	0.36	17.5	−0.073	0.05	1
20	9	36.15-38.36	UMB10000133	0.18	97.5	−0.158	0.24	20
21	9	75.48-76.83	ALGA0053672	0.39	20.6	−0.171	0.16	2
22	9	85.48-85.84	H3GA0027836	0.29	16.4	−0.17	0.07	2
23	10	52.21-53.94	DRGA0010548	0.09	122.3	0.39	0.29	10
24	10	56.37-56.78	ASGA0103067	0.48	35.5	0.101	0.09	2
25	11	25.34-26.23	ALGA0061540	0.3	19.5	0.171	0.05	7
26	12	42.93-44.76	ALGA0123748	0.25	18.5	−0.141	0.14	18
27	12	48.73-49.13	ASGA0082658	0.24	16.4	0.234	0.06	4
28	12	51.95-52.62	ALGA0066876	0.47	156.4	0.094	1.1	14
29	12	54.70-56.06	H3GA0034702	0.19	78.4	0.212	0.29	90
30	14	80.44-84.34	ALGA0079106	0.48	140.4	0.143	0.56	47
31	15	3.02-3.24	H3GA0043638	0.07	17.5	0.309	0.03	1
32	15	8.04-8.91	ASGA0068444	0.26	30.1	0.195	0.06	3
33	15	26.16-26.82	CASI0009989	0.17	16.4	0.235	0.05	1
34	15	147.34-147.74	ASGA0071500	0.25	17.5	0.246	0.04	8
35	15	153.82-154.57	M1GA0027067	0.4	53.1	−0.027	0.35	23
36	16	1.02-1.03	H3GA0045756	0.44	12.3	−0.034	0.04	0
37	16	27.60-30.67	DRGA0016028	0.34	30.1	0.035	0.43	24
38	18	21.22-21.27	H3GA0050517	0.47	13.3	−0.228	0.05	0
39	18	50.12-51.64	ASGA0080142	0.21	15.4	0.237	0.06	6

[1]The position in Mb of the significant (BF >10) left and right flanking markers.
[2]The allele substitution effect is the regression coefficient of the most significant SNP of the QTL on number of teats corrected for fixed effects (sex of the animal and farm). The minor allele is counted.
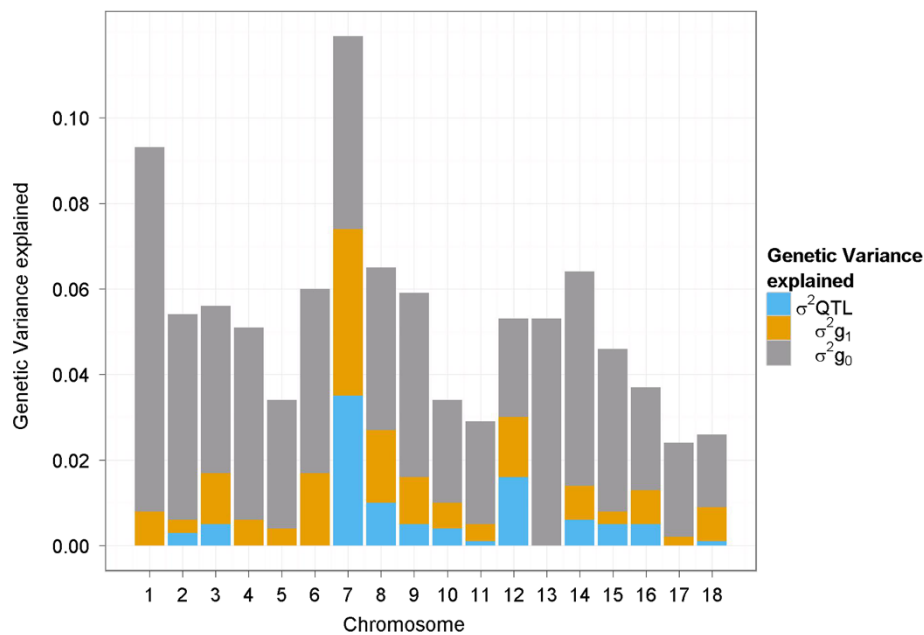[3]The genetic variance explained by the QTL region expressed in %.

Figure 3 Genetic variance explained per chromosome. The variances explained by the defined QTL regions as in Table 1 are shown in blue (σ^2_{QTL}). The variances explained by SNPs when placed in the second distribution (π_1) are shown in orange ($\sigma^2_{g_1}$). The variances explained by SNPs when placed in the null distribution (π_0) are shown in grey ($\sigma^2_{g_0}$). The variances of the SNPs were summed per chromosome.

between 52.21 and 53.94 Mb. Of these, *MPP7*, is located at 53.01 Mb. It encodes a MAGUK (peripheral membrane-associated guanylate kinases) p55 subfamily member 7-like transcript, which is required for the establishment of cell polarity in the developing *Drosophila* embryo [50]. Also in cultured cells, MPP7 promotes epithelial cell polarity and tight junction formation [51]. The second candidate gene is *ARMC4* at position 53.34 Mb. *ARMC4* belongs to the *CTNNB1*-family of beta-catenins, which are well known for roles in transcriptional transduction of *Wnt* signaling and tissue integrity via cell-cell adhesion. The third candidate gene is *MKX* (Mohawk) at 53.68 Mb, a homeobox gene which again acts as an important regulator of vertebrae development. For the two other QTL regions on SSC9 between 36.15 and 38.46 Mb and on SSC14 between 80.44 and 84.34 Mb, no obvious candidate genes were identified.

Discussion

Strength of GWAS methodology

In the present study, a Bayesian Variable Selection approach was used to estimate SNP effects. By using a relatively stringent prior (π_1 =0.001) on average 42 SNPs per run are selected to have an effect on the trait. To secure that many SNPs have been in the π_1 distribution, the number of cycles run was large (500.000 respectively). The advantage is that less SNPs are given a large variance and therefore we expect a clearer

distinction between SNPs with a small effect and SNPs with a larger effect [52]. Instead of a sliding window of N number of SNPs, which is often used to account for linkage disequilibrium (LD) [47], regions were defined based on distances between significant SNPs (<2 Mb) and LD for post-analyses. The number of SNPs within a region was variable between a single SNP up to 16 SNPs. The defined region was used to simultaneously estimate the explained variance of the SNPs within the QTL region.

The use of deregressed EBV's should also give more reliable results from the GWAS resulting in a better estimation of the size of the SNP effects [53]. Use of only phenotypes in a GWAS from the animals genotyped without considering the information of offspring, parents and other family members is reducing the power of the study. Deregressed EBV's are often used in dairy cattle breeding where daughter yield deviation (DYD) are used in GWAS, which have similar properties to deregressed EBV's [54]. Deregressing breeding values is used to circumvent selecting SNPs which explain family relatedness rather than associated genes. This is done by removing the contribution of information from relatives.

Reliability of identified QTL areas

In total 39 QTL regions were identified with relatively small effects, which suggest that NTE is controlled by many genes. This is in agreement with earlier performed linkage studies where across different pig populations

many QTL were found on almost all chromosomes. Including the results from this study and published QTL, all chromosomes with the exception of SSC13 and the Y chromosome, carry QTL for teat number. Figure 3 also clearly shows that some QTL found in this study only have a limited contribution compared to the genetic variance explained by SNPs which were in the null distribution (π_0). This helps to distinguish whether the variance explained per chromosome is expected because SNPs have a small variance, or alternatively, whether there is a large QTL contributing to the chromosomal genetic variance. The expected variance explained per chromosome is proportional to the number of SNPs on the chromosome with the assumption that no QTL is located on the chromosome, for example on SSC13. The difference between the total chromosomal genetic variance and the expected variance when SNPs are in the null distribution plus the QTL variance, can be caused by SNPs that were selected (π_1) but did not reach the significance level to be assigned to a QTL region. The genetic variance explained by the defined QTL region on SSC7 might even be an underestimation due to SNPs which are in LD with the QTL region, but did not reach the significance threshold of a BF >10. In general, the explained variance by the identified QTL regions is small. Several factors could be underlying the relatively small explained variance. The size of this study is moderate in livestock and small compared to human studies [55]. A larger sample size could pick up more rare variants, additionally a larger SNP set (>500.000) will have the SNP closer to the causative mutation and will give more statistical power to the association study [55]. The trait NTE has also been under selection for at least 10 years (E.F. Knol, personal observation), which could have resulted in rapid fixation of large segregating QTL. Additionally, NTE is a categorical trait and not directly measuring the variation of the underlying physiological factors of the developmental signalling chain.

Most of the QTL regions identified in this study reside within published QTL for NTE on several chromosomes (SSC2, 3, 7, 8, 10, 11, 12, 15 and 16 respectively). In Figure 4, a detailed overview of all the published QTL compared to the results from this study are shown (see Additional file 2 for the details of the studies published on NTE). Overall, the regions identified in this study are much smaller, as a result of using high-density SNP information in a GWAS instead of microsatellites as in a linkage study.

Especially on SSC7, 8, 10 and 12 some QTL are located relatively close to each other. We identify them as small individual QTL regions rather than one large QTL, because the LD (measured by r^2) between the SNPs is less than 0.7. Besides region 30 on SSC14, the r^2 between the closely identified QTL was always below 0.2, which suggests the QTL should be considered as independent genomic regions.
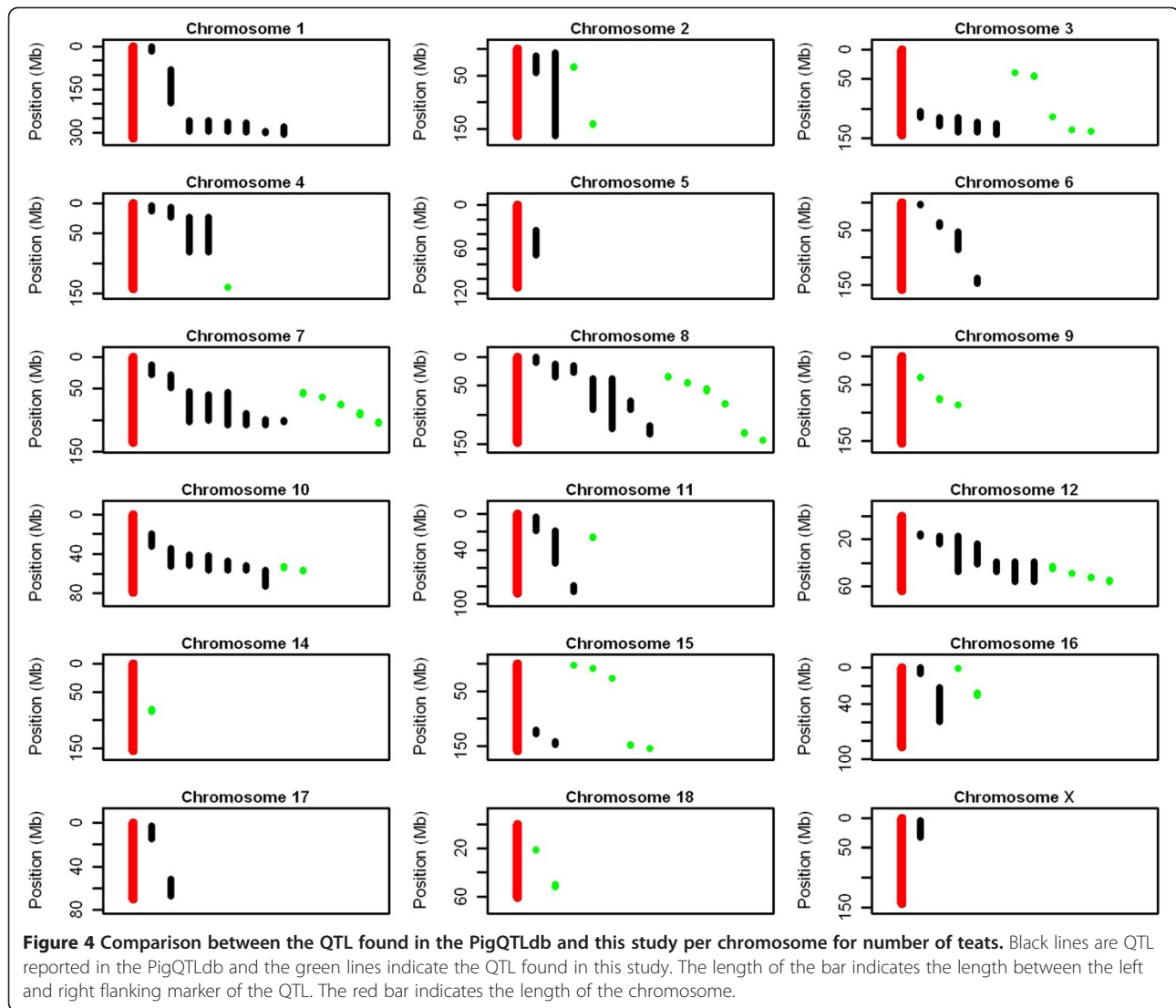
Candidate QTL and genes

The QTL which explains the most genetic variation is located on SSC7 and has been identified in other studies [9,11,12,17,18]. While Mikawa et al. [48] reported an additive effect of this QTL on number of vertebrae, we observed a dominant or even overdominant effect on the number of teats in our population. This is in accordance with a previous study reporting that breeding for increased body length also resulted in increased teat number, leading to the speculation that the number of vertebrae is genetically linked to the number of teats [11]. Within this QTL we identified a gene named *Vertnin* (*VRTN*) at 103.45 Mb, which likely provides such a genetic link. The likeliness of *Vertnin's* candidature is strongly supported by the previous demonstration that structural integrity of the somites, as well as somitic expression of genes such as *Gli3* and *Fgf10*, are required for the proper formation of the mammary line and glands at the axial level of thoracic/abdominal transition in mice [27]. Mechanistically, the genetic linkage between numbers of vertebrae and teats can be explained by the somites being precursors for vertebrae as well as dermal mesenchyme, the latter inducing mammary gland formation in the overlying surface ectoderm [27,39,40].

Another candidate gene in the region is *Prox2*, a vertebrate homolog of the *Drosophila* homeobox-containing gene prospero *Prox2* belongs to a family of transcription factors whose function has not yet been characterized in detail in mammals. In zebrafish, *Prox2* is mainly expressed in the brain and involved in eye development [56]. Notably, both *Vertnin* and *Prox2* have recently been proposed as candidates for the number of vertebrae in a White Duroc x Chinese Erhualian intercross, that also carries two different haplotypes for increased teat number [57].

The second largest QTL effect explaining more than 1% of the genetic variance is located on SSC12 between 51.95 and 52.62 Mb. Although many studies have found a QTL for NTE on SSC12 [9,11,12,17,18], only Guo et al. [12] found a large QTL interval similar to ours in an F2 cross between Large White and Chinese Meishan. Within this region, we find the δ-*EF1* gene is of particular interest because its expression in the somites [58] may suggest a role in vertebrae and mammary development in a mechanism similar to *Vertnin*, as described above.

On SSC10, we found a significant QTL in the same region as in other studies [12,13,15] or neighbouring a region found in other studies [16,59]. Between 52.21 and 53.94 Mb within this region, we identified three

Figure 4 Comparison between the QTL found in the PigQTLdb and this study per chromosome for number of teats. Black lines are QTL reported in the PigQTLdb and the green lines indicate the QTL found in this study. The length of the bar indicates the length between the left and right flanking marker of the QTL. The red bar indicates the length of the chromosome.

candidate genes, namely *MPP7, ARMC4,* and *MKX.* In *Drosophila* and in cultured cells, *MPP7* promotes cell polarity and tight junction formation [50,51]. While a role for *MPP7* and tight junctions in early mammary gland development has not yet been studied, cell polarization and tissue stratification are integral part of mammary induction and early growth, as observed in mice [23], supporting the candidature involvement of *MPP7* establishment of gland/teat number. As a member of the *β-catenin* gene family, *ARMC4* could be involved in mammary gland formation via its role in transcriptional transduction of *Wnt/β-catenin* signaling. A role for this signaling pathway in early mammary gland formation is demonstrated by compromised mammary gland formation in the presence of pathway-inhibitory mutations, and supernumerary mammary gland formation in the context of excessive Wnt signaling [33,34,60-63]. Alternatively, Wnt signaling is implicated downstream of

PTHrP signaling in nipple formation, which is a process that occurs secondary to the formation of mammary rudiments [64].

Most convincing, the third gene *MKX* is in mice expressed in the condensing mesenchyme that will ultimately become the proximal ribs and vertebral bodies [65].

The two other QTL regions on SSC9 between 36.15 and 38.46 Mb and on SSC14 between 80.44 and 84.34 Mb are new and have not been reported before. Soma et al. [66] reported a QTL for number of lumbar vertebrae on SSC14 around 137.88 Mb (119 cM, marker SW761) in a Duroc purebred population, but the QTL peak is located around 55 Mb further distal of our QTL no. 30 (SSC14: 80.44-84.34 Mb).

Elongation of the back and an increased number of vertebrae in pigs as a consequence of domestication was already observed by Charles Darwin [67]. Signatures of selection have also been found for numbers of vertebrae

and body length in the domestic pig [68] and for number of teats [69]. In accordance, some QTL for number of vertebrae and number of teats in pigs overlap [11]. A mutation increasing number of teats in the gene *NR6A1* (an orphan nuclear receptor) on SSC1 has been fixed in commercial breeds [70] and therefore, this QTL is not segregating in our population either. Nevertheless, selection for increased carcass length still provides variation for genetic improvement of reproductive traits such as number of teats in the sow which is in turn very relevant for the survival of piglets.

Mechanistically, the link between carcass length (number of vertebrae) and number of teats can be explained by results from studies in mice, revealing a role for the somites, *i.e.* precursor structures for vertebrae, in the induction of mammary gland development [22,25,27]. Although to our knowledge, a variation in number of somites (vertebrae) in mice has not been subject to study in mice, and certainly not in relationship to the number of mammary glands or nipples/teats, it is clear that altered somitic development or gene expression can alter the number of mammary glands (thus nipples/teats) in mice [22,25,27]. In agreement with this biological mechanism of mammary gland development, we identified in our current study several candidate genes with a known association to vertebrae development. To date, none of these genes have a reported role in mammary gland or nipple/teat development. Such studies could certainly be helpful to get closer to the causative genetic variant which can be used in breeding programs for increased NTE in pigs.

Conclusion
Although NTE has been under selection for many generations, this study found many QTL controlling NTE. We could narrow down considerably some of the earlier published QTL regions making it easier to select candidate genes. Five major QTL were found at high resolution on SSC7, 9, 10, 12, and 14 of which the QTL on SSC9 and SSC14 are the first ones to be reported on these chromosomes. The confirmed major QTL found on SSC7 contains the candidate gene *Vertnin* which has been reported to control the number of thoracic vertebrae. Interestingly, also the two other regions on SSC12 and SSC10 contain genes that have a suspected (δ-*EF1*) or demonstrated (*MKX*) involvement in vertebrae development. The genetic relation between teat number and number of vertebrae can be explained by the somites being precursors for both vertebrae and dermal mesenchyme, while both the somites and dermal mesenchyme have been shown to contain inductive signals for mammary gland formation. All significant QTL together explain almost 10% of the genetic variance. Nevertheless, results clearly show the polygenic nature of

the trait indicating the genetic complexity of the trait. The significant SNPs found in this study could be used in selection to increase NTE in pigs, so that the increasing number of live-born piglets can be nursed by the sow.

Methods
Animals
This study was conducted strictly in line with the regulations of the Dutch law on the protection of animals (Gezondheids- en welzijnswet voor dieren). Animal Care and Use Committee approval was not obtained for this study, because the data were obtained from an existing database. Phenotypic measurements of the number of teats (NTE) were obtained from 936 pigs, of which 230 were boars and 706 sows. All pigs were purebred Large White and were born between 2006 and 2009 and originated from 17 farms. The number of teats were counted at birth and recorded on both sexes. In this study, NTE was the only recorded trait and number of left and right teats and teat malformations were not considered.

The boars originated from 100 sires (29 genotyped) and 148 dams (20 genotyped). The sows originated from 164 sires (31 genotyped) and 464 dams (48 genotyped). The boars had between 0 to 3,052 offspring with phenotypic observations and the sows between 0 and 180 offspring with phenotypic observations.

Genotyping and SNP quality
Genotyping was performed using the Illumina PorcineSNP60 Beadchip. Samples collected for DNA extraction were only used for routine diagnostic purpose of the breeding program and was strictly in line with the Dutch law on the protection of animals (Gezondheids- en welzijnswet voor dieren). DNA was extracted from blood, hair and ear punches and commercially genotyped at Service XS (Leiden, The Netherlands) or at Geneseek, Inc. (Lincoln, NE, USA). SNPs with a GenCall score <0.7, call rate <95%, minor allele frequency <0.01 and SNPs with no physical position on the pig genome (pig genome build10.2) were removed. After these quality control measures, 42,654 SNPs out of 64,232 SNPs remained for the genome-wide association (GWA).

Statistical method for GWA analyses
The estimated breeding value (EBV) for every animal was obtained via routine genetic evaluation using MiXBLUP [71] in a multitrait model. The model for obtaining the EBV for NTE included fixed effects for herd-year-season, sex and line and an additive genetic effect (animal) as a random effect. Reliabilities per animal were extracted from the genetic evaluation and was based on the methodology of Tier and Meyer [72]. The EBV was deregressed using the methodology proposed by Garrick et al. [47].

The deregressed EBV only contains information of the animals' own performance and of their descendants', which was achieved by removing the parent average. The reliabilities (information sources) vary considerably between animals and therefore the deregressed EBVs have heterogeneous variances. This is resolved by weighing the residuals as in Garrick et al. [47]. Deregression of the EBV was applied to account for the dense family structure in the data and the large difference in the number of information sources to avoid double counting.

A Bayesian Variable Selection model [73] was fitted for NTE by estimating the marker effects with all SNPs simultaneously in the model:

$$y = \mu + X\beta + e,$$

where **y** is a n-vector of phenotypes on n animals, μ is a n-vector equal to the mean, X is a n by p matrix where p SNPs are coded as 0, 1, or 2 copies of a particular allele vector and β is a p-vector with the marker effects. A Bernoulli distribution is applied on the marker effect:

$$\beta \sim \begin{cases} N\left(0, \sigma_{g_0}^2\right) & \text{with probability: } \pi_0 \\ N\left(0, \sigma_{g_1}^2\right) & \text{with probability: } \pi_1 = 1 - \pi_0 \end{cases}$$

where the first distribution is referred to the null distribution and SNPs are assumed to have a small effects ($\sigma_{g_0}^2$) and the second distribution contains SNPs assumed to have a large effect explaining a large variance ($\sigma_{g_1}^2$) of the phenotype. The probability to be in the null distribution (π_0) was set to 0.999, meaning only 1 in every 1000 SNPs will be in the second distribution which is on average 42 SNP per cycle. The term e is a n-vector with random residual effects assumed to be normally distributed but weighted, $N\left(0, \sigma_e^2 W\right)$, where W is a diagonal matrix with elements $w_1, ..., w_n$. The model was implemented in the program Bayz [74].

A total of 500,000 MCMC chains with a burn-in of 5,000 cycles were run and a Metropolis-Hastings sampler was applied to get good convergence which was assessed by visual inspection of the trace and using Gelman and Rubin's convergence diagnostic based on deviance [75] using the R package CODA [76].

Identification of associated regions

To determine which SNPs are significantly associated, a Bayes Factor was calculated for every SNP using the prior probability (π_0 and π_1) and the posterior probability ($\hat{p}_i$) by calculating an odds ratio as:

$$BF = \frac{\hat{p}_i / (1 - \hat{p}_i)}{\pi_1 / \pi_0}$$

A BF >10 is referred to as 'strong' and a value above 100 as 'decisive' [77].

When at least two SNPs in a region (<2 Mb) showed a BF >10, this region was defined as a candidate QTL region and for regions with a BF around or over 100, a gene search was conducted. To define a QTL for NTE in this analysis, linkage disequilibrium (LD) was taken into account. When r^2 was >0.7 between two QTL regions, but the distance was larger than 2 Mb, the region was still combined into one common region.

The variance explained by the QTL region was calculated by simultaneously estimating the variance explained by all the regions and all the other non-significant SNPs. To get insight in the chromosomal partitioning of the genetic variance, SNPs within a QTL region (BF > 10) and SNPs on different chromosomes were placed in different groups. The sum of the variance explained per chromosome was the sum of variances of the QTL (if detected) on a chromosome plus the explained variance by the other SNPs on the chromosome. The variance expected per chromosome was calculated as the average variance explained by SNPs when placed in the null distribution (π_0).

QTL comparisons and candidate genes

All earlier reported QTL found for NTE were available at the PigQTLdb (http://www.animalgenome.org/QTLdb/pig.html). Flanking markers of the QTL were searched at the reference genome (build10.2) to find the physical position of the markers. If one of the markers was not found, a BLAST search was performed. If the physical position of markers could not be identified the closest neighboring marker according to the linkage map from MARC USDA (http://www.marc.usda.gov/genome) was mapped. When any of these markers could not be placed on the physical map, the QTL was not included in the comparison.

For gene searches, the left and right flanking markers of the defined candidate QTL regions were used. Pig genome build10.2 was used for the position of the SNPs. Gene annotation for QTL regions was performed with BIOMART software in the Ensembl Sscrofa 10.2 (http://www.ensembl.org). Ensemble Gene IDs were used to count the number of genes within the QTL regions.

Additional files

Additional file 1: All 489 genes found in QTL regions as reported in Table 1. The genes are mapped on build10.2. Ensemble gene id was given, start and end of the gene, the status of the transcript and gene and when known the gene name and function.

Additional file 2: Overview of the studies reporting QTL on number of teats. The position of the microsatellites are given in Mb and were mapped on build10.2 [9-18,59,78-85].

Competing interest

or preparation of the manuscript. The authors declare that they have no competing interests.

Authors' contributions
ND was involved in conducting the statistical analyses, prepared figures and tables, and drafted the paper. JMV was a discussion partner and involved in writing the paper. BH was involved in the selection of the candidate genes and writing the paper. All authors have read and approved the final manuscript.

Author details
[1]TOPIGS Research Center IPG, PO Box 43, 6640AA Beuningen, The Netherlands. [2]Institute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), 61, Biopolis Drive, Singapore, Singapore 138673.

References
1. Merks JWM, Mathur PK, Knol EF: **New phenotypes for new breeding goals in pigs.** *Animal* 2012, **6**:535–543.
2. Neeteson-van Nieuwenhoven A-M, Knap P, Avendaño S: **The role of sustainable commercial pig and poultry breeding for food security.** *Animal Frontiers* 2013, **3**:52–57.
3. Bell AG: **On the development by selection of supernumerary mammae in sheep.** *Science* 1898, **9**:6337–6339.
4. Schmidt H: **Supernumerary nipples: prevalence, size, sex and side predilection -a prospective clinical study.** *Eur J Pediatr* 1998, **157**:821–823.
5. Chalkias H, Rydhmer L, Lundeheim N: **Genetic analysis of functional and non-functional teats in a population of Yorkshirepigs.** *Livest Sci* 2013, **152**:127–134.
6. McKay RM, Rahnefeld GW: **Heritability of teat number in swine.** *Can J Anim Sci* 1990, **70**:425–430.
7. Hu ZL, Park CA, Wu XL, Reecy JM: **Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era.** *Nucleic Acids Res* 2013, **41**:D871–D879.
8. Hu ZL, Fritz ER, Reecy JM: **AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond.** *Nucleic Acids Res* 2007, **35**:D604–D609.
9. Bidanel JP, Rosendo A, Iannuccelli N, Riquet J, Gilbert H, Caritez JC, Billon Y, Amigues Y, Prunier A, Milan D: **Detection of quantitative trait loci for teat number and female reproductive traits in Meishan X Large White F2 pigs.** *Animal* 2008, **2**:813.
10. Cassady JP, Johnson RK, Pomp D, Rohrer GA, Van Vleck LD, Spiegel EK, Gilson KM: **Identification of quantitative trait loci affecting reproduction in pigs.** *J Anim Sci* 2001, **79**:623–633.
11. Ding N, Guo Y, Knorr C, Ma J, Mao H, Lan L, Xiao S, Ai H, Haley CS, Brenig B, Huang L: **Genome-wide QTL mapping for three traits related to teat number in a White Duroc x Erhualian pig resource population.** *BMC Genet* 2009, **10**:6.
12. Guo YM, Lee GJ, Archibald AL, Haley CS: **Quantitative trait loci for production traits in pigs: a combined analysis of two Meishan x Large White populations.** *Anim Genet* 2008, **39**:486–495.
13. Hirooka H, De Koning DJ, Harlizius B, Van Arendonk JA, Rattink AP, Groenen MA, Brascamp EW, Bovenhuis H: **A whole-genome scan for quantitative trait loci affecting teat number in pigs.** *J Anim Sci* 2001, **79**:2320–2326.
14. Holl JW, Cassady JP, Pomp D, Johnson RK: **A genome scan for quantitative trait loci and imprinted regions affecting reproduction in pigs.** *J Anim Sci* 2004, **82**:3421–3429.
15. Rodriguez C, Tomas A, Alves E, Ramirez O, Arque M, Munoz G, Barragan C, Varona L, Silio L, Amills M, Noguera JL: **QTL mapping for teat number in an Iberian x Meishan pig intercross.** *Anim Genet* 2005, **36**:490–496.
16. Rohrer GA: **Identification of quantitative trait loci affecting birth characters and accumulation of backfat and weight in a Meishan-White Composite resource population.** *J Anim Sci* 2000, **78**:2547–2553.
17. Sato S, Atsuji K, Saito N, Okitsu M, Komatsuda A, Mitsuhashi T, Nirasawa K, Hayashi T, Sugimoto Y, Kobayashi E: **Identification of quantitative trait loci affecting corpora lutea and number of teats in a Meishan x Duroc F2 resource population.** *J Anim Sci* 2006, **84**:2895–2901.
18. Zhang J, Xiong Y, Zuo B, Lei M, Jiang S, Li F, Zheng R, Li J, Xu D: **Detection of quantitative trait loci associated with several internal organ traits and teat number trait in a pig population.** *J Genet Genomics* 2007, **34**:307–314.
19. Propper AY, Howard BA, Veltmaat JM: **Prenatal morphogenesis of mammary glands in mouse and rabbit.** *J Mammary Gland Biol Neoplasia* 2013. in press.
20. Veltmaat JM, Mailleux AA, Thiery JP, Bellusci S: **Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation.** *Differentiation* 2003, **71**:1–17.
21. Dixon JM: *ABC of Breast Diseases.* Chichester: John Wiley & Sons; 2012.
22. Veltmaat JM, Van Veelen W, Thiery JP, Bellusci S: **Identification of the mammary line in mouse by Wnt10b expression.** *Dev Dynam* 2004, **229**:349–356.
23. Lee MY, Racine V, Jagadpramana P, Sun L, Yu W, Du T, Spencer-Dene B, Rubin N, Le L, Ndiaye D, Bellusci S, Kratochwil K, Veltmaat JM: **Ectodermal Influx and Cell Hypertrophy Provide Early Growth for All Murine Mammary Rudiments, and Are Differentially Regulated among Them by Gli3.** *PLoS One* 2011, **6**:e26242.
24. Propper AY: **Wandering epithelial cells in the rabbit embryo milk line: a preliminary scanning electron microscope study.** *Dev Biol* 1978, **67**:225–231.
25. Veltmaat JM, Ramsdell AF, Sterneck E: **Positional Variations in Mammary Gland Development and Cancer.** *J Mammary Gland Biol Neoplasia* 2013, **18**:1–10.
26. Howard B, Panchal H, McCarthy A, Ashworth A: **Identification of the scaramanga gene implicates Neuregulin3 in mammary gland specification.** *Gene Dev* 2005, **19**:2078–2090.
27. Veltmaat JM, Fdr R, Le LT, Kratochwil K, Sala FG, Van Veelen W, Rice R, Spencer-Dene B, Mailleux AA, Rice DP, Thiery JP, Bellusci S: **Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes.** *Development* 2006, **133**:2325–2335.
28. Mailleux AA, Spencer-Dene B, Dillon C, Ndiaye D, Savona-Baron C, Itoh N, Kato S, Dickson C, Thiery JP, Bellusci S: **Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo.** *Development* 2002, **129**:53–60.
29. Garcia-Gasca A, Spyropoulos DD: **Differential mammary morphogenesis along the anteroposterior axis in Hoxc6 gene targeted mice.** *Dev Dynam* 2000, **219**:261–276.
30. Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR, Bradley A: **p63 is a p53 homologue required for limb and epidermal morphogenesis.** *Nature* 1999, **398**:708–713.
31. Yang A, Schweitzer R, Sun D, Kaghad M, Walker N, Bronson RT, Tabin C, Sharpe A, Caput D, Crum C, McKeon F: **p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development.** *Nature* 1999, **398**:714–718.
32. Eblaghie MC, Song S, Kim J, Akita K, Tickle C, Jung H: **Interactions between FGF and Wnt signals and Tbx3 gene expression in mammary gland initiation in mouse embryos.** *J Anat* 2004, **205**:1–13.
33. van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L, Grosschedl R: **Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice.** *Gene Dev* 1994, **8**:2691–2703.
34. Gu B, Sun P, Yuan Y, Moraes RC, Li A, Teng A, Agrawal A, Rheaume C, Bilanchone V, Veltmaat JM, Takemaru KI, Millar SE, Lee EYHP, Lewis MT, Li B, Dai X: **Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation.** *J Cell Biol* 2009, **185**:811–826.
35. Foley J, Dann P, Hong J, Cosgrove J, Dreyer B, Rimm D, Dunbar M, Philbrick W, Wysolmerski JJ: **Parathyroid hormone-related protein maintains mammary epithelial fate and triggers nipple skin differentiation during embryonic breast development.** *Development* 2001, **128**:513–525.
36. Mustonen T, Ilmonen M, Pummila M, Kangas AT, Laurikkala J, Jaatinen R, Pispa J, Gaide O, Schneider P, Thesleff I, Mikkola ML: **Ectodysplasin A1 promotes placodal cell fate during early morphogenesis of ectodermal appendages.** *Development* 2004, **131**:4907–4919.
37. Panchal H, Wansbury O, Parry S, Ashworth A, Howard B: **Neuregulin3 alters cell fate in the epidermis and mammary gland.** *BMC Dev Biol* 2007, **7**:105.
38. Ahn Y, Sims C, Logue JM, Weatherbee SD, Krumlauf R: **Lrp4 and Wise interplay controls the formation and patterning of mammary and other

skin appendage placodes by modulating Wnt signaling. *Development* 2013, **140**:583–593.

39. Närhi K, Tummers M, Ahtiainen L, Itoh N, Thesleff I, Mikkola ML: **Sostdc1 defines the size and number of skin appendage placodes.** *Dev Biol* 2012, **364**:149–161.

40. Propper A, Gomot L: **Tissue interactions during organogenesis of the mammary gland in the rabbit embryo.** *C R Acad Sci Hebd Seances Acad Sci D* 1967, **264**:2573–2575.

41. Kratochwil K: **Organ specificity in mesenchymal induction demonstrated in the embryonic development of the mammary gland of the mouse.** *Dev Biol* 1969, **20**:46–71.

42. Cho KW, Kwon HJ, Shin JO, Lee JM, Cho SW, Tickle C, Jung HS: **Retinoic acid signaling and the initiation of mammary gland development.** *Dev Biol* 2012, **365**:259–266.

43. Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P, Hansen MS, Hedegaard J, Hu ZL, Kerstens HH, Law AS, Megens HJ, Milan D, Nonneman DJ, Rohrer GA, Rothschild MF, Smith TPL, Schnabel RD, Van Tassel CP, Taylor JF, Wiedmann RT, Schook LB, Groenen MAM: **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.

44. Duijvesteijn N, Knol EF, Merks JWM, Crooijmans RPMA, Groenen MAM, Bovenhuis H, Harlizius B: **A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6.** *BMC Genet* 2010, **11**:42.

45. Fan B, Onteru SK, Du ZQ, 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.

46. 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**:988–995.

47. Garrick DJ, Taylor JF, Fernando RL: **Deregressing estimated breeding values and weighting information for genomic regression analyses.** *Genet Sel Evol* 2009, **41**:55.

48. Mikawa S, Sato S, Nii M, Morozumi T, Yoshioka G, Imaeda N, Yamagushi T, Awata T: **Identification of a second gene associated with variation in vertebral number in domestic pigs.** *BMC Genet* 2011, **12**:5.

49. Bellon E, Luyten FP, Tylzanowski P: **d-EF1 is a negative regulator of Ihh in the developing growth plate.** *J Cell Biol* 2009, **187**:685–699.

50. Bohl J, Brimer N, Lyons C, Pol SBV: **The stardust family protein MPP7 forms a tripartite complex with LIN7 and DLG1 that regulates the stability and localization of DLG1 to cell junctions.** *J Bio Chem* 2007, **282**:9392–9400.

51. Stucke VM, Timmerman E, Vandekerckhove J, Gevaert K, Hall A: **The MAGUK protein MPP7 binds to the polarity protein hDlg1 and facilitates epithelial tight junction formation.** *Mol Biol Cell* 2007, **18**:1744–1755.

52. Van den Berg I, Fritz S, Boichard D: **QTL fine mapping with Bayes C(π): a simulation study.** *Genet Sel Evol* 2013, **45**:19.

53. Ostersen T, Christensen OF, Henryon M, Nielsen B, Su G, Madsen P: **Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in pure-bred pigs.** *Genet Sel Evol* 2011, **43**:38.

54. Bolormaa S, Pryce JE, Hayes BJ, Goddard ME: **Multivariate analysis of a genome-wide association study in dairy cattle.** *J Dairy Sci* 2010, **93**:3818–3833.

55. Visscher PM, Brown MA, McCarthy MI, Yang J: **Five years of GWAS discovery.** *Am J Hum Genet* 2012, **90**:7–24.

56. Pistocchi A, Bartesaghi S, Cotelli F, Del Giacco L: **Identification and expression pattern of zebrafish prox2 during embryonic development.** *Dev Dynam* 2008, **237**:3916–3920.

57. Ren DR, Ren J, Ruan GF, Guo YM, Wu LH, Yang GC, Zhou LH, Li L, Zhang ZY, Huang LS: **Mapping and fine mapping of quantitative trait loci for the number of vertebrae in a White Duroc x Chinese Erhualian intercross resource population.** *Anim Genet* 2012, **43**:545–551.

58. Ji F, Sekido R, Murai K, Kamachi Y, Kondoh H: **d-crystallin enhancer binding protein dEF1 is a zinc finger-homeodomain protein implicated in postgastrulation embryogenesis.** *Development* 1993, **119**:433–446.

59. Dragos-Wendrich M, Moser G, Bartenschlager H, Reiner G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 10.** *J Anim Breed Genet* 2003, **120**:82–88.

60. Boras-Granic K, Chang H, Grosschedl R, Hamel PA: **Lef1 is required for the transition of Wnt signaling from mesenchymal to epithelial cells in the mouse embryonic mammary gland.** *Dev Biol* 2006, **295**:219–231.

61. Chu EY, Hens J, Andl T, Kairo A, Yamaguchi TP, Brisken C, Glick A, Wysolmerski JJ, Millar SE: **Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis.** *Development* 2004, **131**:4819–4829.

62. Lindvall C, Evans NC, Zylstra CR, Li Y, Alexander CM, Williams BO: **The Wnt signaling receptor Lrp5 is required for mammary ductal stem cell activity and Wnt1-induced tumorigenesis.** *J Biol Chem* 2006, **281**:35081–35087.

63. Lindvall C, Zylstra CR, Evans N, West RA, Dykema K, Furge KA, Williams BO: **The Wnt co-receptor Lrp6 is required for normal mouse mammary gland development.** *PLoS One* 2009, **4**:e5813.

64. Hiremath M, Dann P, Fischer J, Butterworth D, Boras-Granic K, Hens J, Van Houten J, Shi W, Wysolmerski JJ: **Parathyroid hormone-related protein activates Wnt signaling to specify the embryonic mammary mesenchyme.** *Development* 2012, **139**:4239–4249.

65. Anderson DM, Arredondo J, Hahn K, Valente G, Martin JF, Wilson-Rawls J, Rawls A: **Mohawk is a novel homeobox gene expressed in the developing mouse embryo.** *Dev Dynam* 2006, **235**:792–801.

66. Soma Y, Uemoto Y, Sato S, Shibata T, Kadowaki H, Kobayashi E, Suzuki K: **Genome-wide mapping and identification of new quantitative trait loci affecting meat production, meat quality, and carcass traits within a Duroc purebred population.** *J Anim Sci* 2011, **89**:601–608.

67. Darwin C: *The Variation of Animals and Plants under Domestication.* 2nd edition. Murray; 1868.

68. Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg O, Jern P, Jorgensen CB, Archibald AL, Fredholm M, Groenen MAM, Anderson L: **Strong signatures of selection in the domestic pig genome.** *Proc Natl Acad Sci U S A* 2012, **109**:19529–19536.

69. Wilkinson S, Lu ZH, Megens H-J, Archibald AL, Haley C, Jackson IJ, Groenen MAM, Crooijmans RPMA, Ogden R, Wiener P: **Signatures of diversifying selection in European pig breeds.** *PLoS Genet* 2013, **9**:e1003453.

70. Mikawa S, Morozumi T, Shimanuki SI, Hayashi T, Uenishi H, Domukai M, Okumura N, Awata T: **Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1).** *Genome Res* 2007, **17**:586–593.

71. Mulder HA, Lidauer M, Strandén I, Mäntysaari EA, Pool MH, Veerkamp RF: *MiXBLUP Manual.* Lelystad, the Netherlands: Animal Breeding and Genomics Centre, Wageningen UR Livestock Research; 2012.

72. Tier B, Meyer K: **Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models.** *J Anim Breed Genet* 2004, **121**:77–89.

73. George EI, McCulloch RE: **Variable selection via Gibbs sampling.** *J Am Stat Assoc* 1993, **88**:881–889.

74. Janss L: **BayZ manual (available at http://bayz.biz/).** 2011.

75. Gelman A, Rubin DB: **Inference from iterative simulation using multiple sequences.** *Stat Sci* 1992, **7**:457–472.

76. Plummer M, Best N, Cowles K, Vines K: **CODA: Convergence diagnosis and output analysis for MCMC.** *R news* 2006, **6**:7–11.

77. Kass RE, Raftery AE: **Bayes factors.** *J Am Stat Assoc* 1995, **90**:773–795.

78. King AH, Jiang Z, Gibson JP, Haley CS, Archibald AL: **Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8.** *Biol Reprod* 2003, **68**:2172–2179.

79. Wada Y, Akita T, Awata T, Furukawa T, Sugai N, Ishii K, Ito Y, Kobayashi E, Mikawa S, Yasue H, Sugai N, Inage Y, Kusumoto H, Matsumoto T, Miyake M, Murase A, Shimanuki S, Sugiyama T, Uchida Y, Yanai S: **Quantitative trait loci (QTL) analysis in a Meishan X Göttingen cross population.** *Anim Genet* 2000, **31**:376–384.

80. Beeckmann P, Moser G, Bartenschlager H, Reiner G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 8.** *J Anim Breed Genet* 2003, **120**:66–73.

81. Lee SS, Chen Y, Moran C, Cepica S, Reiner G, Bartenschlager H, Moser G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 2.** *J Anim Breed Genet* 2003, **120**:11–19.

82. Lee SS, Chen Y, Moran C, Stratil A, Reiner G, Bartenschlager H, Moser G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 5.** *J Anim Breed Genet* 2003, **120**:38–44.

83. Cepica S, Reiner G, Bartenschlager H, Moser G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome X.** *J Anim Breed Genet* 2003, **120**:144–151.
84. Beeckmann P, Schröffel J, Moser G, Bartenschlager H, Reiner G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 1.** *J Anim Breed Genet* 2003, **120**:1–10.
85. Yue G, Schröffel J, Moser G, Bartenschlager H, Reiner G, Geldermann H: **Linkage and QTL mapping for Sus scrofa chromosome 12.** *J Anim Breed Genet* 2003, **120**:95–102.