451. Allele frequency differences at epistatic QTL explain different genetic trends in number of teats in two pig lines

C.A. Sevillano[1*], B. Harlizius[1], M.F.L. Derks[1,2], M.S. Lopes[1,3], M. van Son[4] and E.F. Knol[1]

[1]Topigs Norsvin Research Center, P.O. Box 43, 6640 AA Beuningen, the Netherlands; [2]Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, the Netherlands; [3]Topigs Norsvin, Visconde do Rio Branco 1310, 80.420-210 Curitiba, Brazil; [4]Norsvin, Storhamargata 44, 2317 Hamar, Norway; claudia.sevillanodelaguila@topigsnorsvin.com

Abstract

Several QTL regions affecting number of teats have been detected in commercial pig lines. In this study we follow the indirect effects of index selection on different QTL regions for number of teats in two maternal lines in 40K animals with imputed 555K SNP. In total, 3 QTL regions overlap between the two populations. For a QTL on *Sus scrofa* chromosome 7 (*SSC7*), the underlying functional variation affecting number of thoracic vertebrae located in the vertnin gene has also been genotyped showing an allelic substitution effect of nearly 0.4 teats in both lines. However, allele frequencies show an opposite trend at the vertnin gene locus in the two lines. Moreover, epistatic effects between two QTL regions on *SSC7* and *SSC12* are investigated at the molecular and phenotypic level.

Introduction

Some mammals, such as mice and pigs, show variation in the number of mammary glands. In pigs, number of teats (NTE) are monitored in commercial breeding and are part of the overall selection index. The aim of this work is to identify how the selection index indirectly affects QTL regions for NTE. In a previous GWAS (van Son *et al.* 2019), we showed in different lines the segregation of the two variants at the vertnin *(VRTN)* locus which is known to cause the development of additional thoracic vertebrae and ribs. Every rib has the potential to develop an additional mammary gland and therefore also explains the effect of this QTL on NTE (Veltmaat, 2017). In addition, several other QTL regions for NTE have been reported in pigs for which the molecular basis remains unknown (Moscatelli *et al.,* 2020, Zhuang *et al.,* 2020). In this study, we identified two additional QTL in our maternal lines on *SSC10* and *SSC12.* At all three QTL, allele frequencies change over time, however, for some loci allele frequencies change in the opposite direction in the two lines. Moreover, the allelic effects of the second largest QTL located at *SSC12* differs between homozygous genotypes at the *VRTN* locus on *SSC7.* This difference can be due to distinct interactions of the *VRTN* locus with different local genetic background, i.e. epistasis (Mackay, 2014).

Materials & methods

Phenotypes, genotypes and pre-processing. We evaluated NTE in two pig populations L (Landrace) and LW (Large-White based). NTE was recorded at birth on both males and females. The L population animals were born between 2013 and 2020. The LW population animals were born between 2009 and 2020.

Two datasets from each population were used: ALL and GENOTYPED. The dataset ALL consisted of all genotyped animals and their contemporaries that had phenotypes (around 300,000 animals per line). Using ALL, the phenotypes were pre-corrected for all non-genetic effects (sex, heard-year, and litter). The non-genetic effects were estimated with a pedigree-based linear model in ASReml v3.0 (Gilmour *et al.*, 2009). The dataset GENOTYPED was a subset of ALL consisting of all animals that had both phenotypes and genotypes (39,956 L and 39,945 LW). This dataset was used to perform the GWAS. All GENOTYPED animals were genotyped on Illumina GeneSeek custom SNP chips (25K, 50K or 80K) (Lincoln, NE, USA). Genotypes of the animals were imputed within population using Fimpute v3 (Sargolzaei *et al.*, 2014) to

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_451, © C.A. Sevillano *et al.* 2022

a set of 48,075 SNPs. In addition, some boars were genotyped using the Axiom porcine 660K array from Affymetrix (Affymetrix Inc., Santa Clara, CA, United States). A second imputation was performed within population using Fimpute v3 (Sargolzaei *et al.*, 2014) from the 48,075 SNPs set to a set of 554,768 SNPs. Quality control was performed at each level of imputation as described in Van Son *et al.* (2019).

Genome-wide association analysis. A single-SNP GWAS was performed with the GENOTYPED dataset within each population using a linear animal model in GCTA software (Yang *et al.*, 2011). Only SNPs with a MAF≥0.01 were used in this step. Significant SNPs and QTL were detected using a $P<1.0\times10^{-8}$.

Results

The mean value for NTE is around 1 teat higher in the L compared to LW (16.72 vs 15.85 for GENOTYPED animals). The L line has gained 1 extra teat in the last 8 years, while the LW line has only gained 0.6 extra teat in the last 10 years (Figure 1). The heritability (h^2) was 0.39 for L and 0.34 for LW.

The GWAS results for NTE in both lines clearly show the QTL on *SSC7* which has been reported earlier to be associated with NTE (van Son *et al.* 2019). In addition, 5 other QTLs segregate on several other chromosomes in the L line, and another 4 QTLs in the LW line. Only 3 QTL regions overlap between the two populations, including the QTL on *SSC7* and the QTL regions on *SSC10* and *SSC12*. The position of the most significant SNP (top SNP) differs slightly between populations for these 3 QTL regions. On *SSC7*, the *VRTN* promoter SNP (VRTN_RS709317845) is the most significant in LW, while the SNP AX-116329721 is the top SNP in L. However, the *VRTN* promoter SNP and the AX-116329721 SNP are in high LD (r^2) (0.98 in L and 0.91 in LW). On *SSC10*, the top SNPs for each line are 124 kb apart and they show low LD in both populations (0.004 in L and 0.002 in LW). On *SSC12*, the top SNPs for each line are 96 kb apart and they are in LD in L population (0.67) but not in LW population (0.002).

In Table 1, we show the parameters for the top SNPs in the QTL regions of *SSC7* and *SSC12*, as they have the largest effect for NTE. For *SSC7,* we limited to the *VRTN* promoter SNP as it is in high LD with the other top SNP in the same region. The top SNP in *SSC12* for line L (AX-116444814) is not present in the GWAS results in LW line due to MAF lower than 0.01. The frequency of the allele C of the *VRTN* promoter SNP increasing NTE is more than four times higher in the L line compared to LW (0.75 vs 0.17) (Table 1). Figure 2 shows that the frequency has been increased over the years in L line, whereas it is dropping in the LW

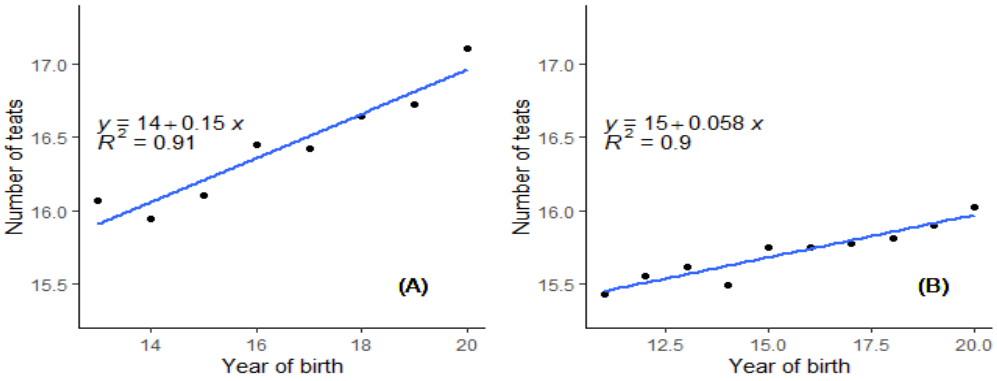

Figure 1. Phenotypic trend for number of teats. (A) Landrace population from 2013 to 2020. (B) Large White population from 2011 to 2020.

Table 1. Parameters for the SNPs in the 2 top QTL associated with NTE in each line.

Line[1]	SSC	SNP	Position (Mb)	*P*-value	Freq.[2]	Effect[2]	SD[2]
L	7	VRTN_RS709317845	98	3.05e-58	0.75	0.37	0.02
	12	AX-116444814	51	4.97e-37	0.17	0.29	0.02
	12	AX-116758961	51	2.68e-10	0.22	0.13	0.02
LW	7	VRTN_RS709317845	98	3.11e-97	0.17	0.38	0.02
	12	AX-116444814	51	NA	0.01	NA	NA
	12	AX-116758961	51	2.73e-41	0.70	0.16	0.01

[1] Line Landrace (L) and Large White (LW).
[2] Relative to the allele increasing NTE.

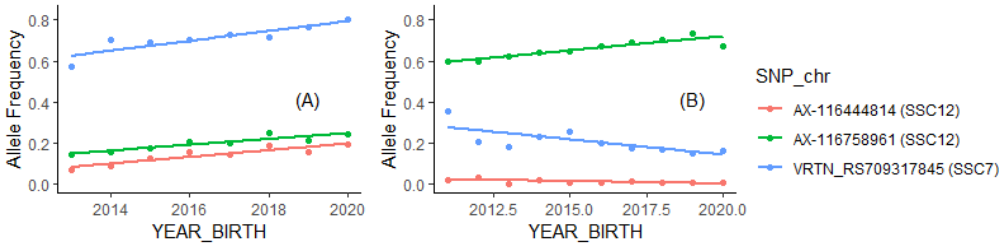

Figure 2. Allele frequency trend for three SNPs associated with number of teats. (A) Landrace population from 2013 to 2020. (B) Large White population from 2011 to 2020.

line. At the same time, the allele increasing NTE on *SSC12* is slightly increasing in both lines. The allele of the top SNP increasing NTE on *SSC12* for LW is much more frequent in LW compared to L (0.22 vs 0.70).

Table 2 shows that allele substitution effects at the *SSC12* QTL are dependent on genotype at the *VRTN* locus on SSC7 and differ between lines. In general, the size of the effect is much higher in *VRTN*-CC animals than in *VRTN*-AA animals. The top SNP in LW (AX-116758961) at the QTL on *SSC12* has a much larger effect in LW than in L but only in *VRTN*-AA animals (0.14 vs 0.06). At the same time, the top SNP in L (AX-116444814) at the QTL on *SSC12* has a lower effect in LW than in L, especially for *VRTN*-AA animals where the effect is even negative (-0.07).

Discussion

This study shows that selection acts on different QTL for NTE in the two sow lines although three QTLs overlaps across lines. However, epistatic effects seems to be playing a role between the two most significant QTL on *SSC7* and *SSC12* that overlap across the two lines. In both lines, CC animals at the *VRTN* locus have a higher genetic potential to develop additional teats than AA animals. Furthermore, the allele frequencies and their trend are the opposite in both lines. At the *VRTN* locus, the allele increasing NTE is going to fixation in L and rather low and declining in LW. At the *SSC12* QTL, the frequency of the allele increasing NTE is very high (0.70) in LW while it is only 0.20 in L. In addition, the top SNP in both lines are different and not in LD in LW. Altogether, this results in different selection pressure at the two QTL. A large proportion of the L animals carry the C allele increasing NTE whereas more than 50% of the LW animals are homozygous for the A allele. Therefore, the unknown causal mutation at the QTL on *SSC12* contributes more to the overall genetic potential and selection space for NTE in LW. Finally, the allele increasing NTE at the top SNP of the *SSC12* QTL in L is hardly present in LW and does not increase NTE in *VRTN*-AA

Table 2. Allelic effects of the largest QTL on SSC12 according to genotypes at the *VRTN* locus on *SSC7*.

Line	SNP	*VRTN*		N of animals[1]	
		CC	AA	CC	AA
L	AX-116444814	0.32	0.12	22,569	2,521
	AX-116758961	0.15	0.06		
LW	AX-116444814	0.27	-0.07	1,245	27,358
	AX-116758961	0.16	0.14		

[1] Number of animals per *VRTN* genotype.

animals. In any case, we clearly show an increased genetic potential to develop additional teats in *VRTN*-CC animals. Altogether, this example shows that rare variants are neglected by genomic selection and could get more emphasis by marker-assisted selection approaches in the breeding program. Additional analyses including data on number of vertebrae and ribs, as well as detailed sequence data, should help to pinpoint the molecular basis of this increased selection space of the *VRTN*-C allele for NTE.

References

Gilmour, A. R., Cullis, B. R., Gogel, B. J., Welham, S. J., and Thompson, R. (2009). ASReml User Guide Release 3.0. VSN International Ltd, England.

Mackay TF. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nat Rev Genet. 2014;15:22. https://doi.org/10.1038/nrg3627

Moscatelli G, Dall'Olio S, Bovo S, Schiavo G, Kazemi H, *et al.* Genome-wide association studies for the number of teats and teat asymmetry patterns in Large White pigs. (2020) Animal genetics 51(4):595-600. https://doi.org/10.1111/age.12947

Sargolzaei, M., Chesnais, J. P., and Schenkel, F. S. A new approach for efficient genotype imputation using information from relatives. (2014) BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478

Van Son M, Lopes MS, Martell HJ, Derks MF, Gangsei LE, *et al.* A QTL for number of teats shows breed specific effects on number of vertebrae in pigs: bridging the gap between molecular and quantitative genetics. (2019) Frontiers in genetics 10:272. https://doi.org/10.3389/fgene.2019.00272

Veltmaat, J. M. (2017). Prenatal Mammary Gland Development in the Mouse: Research Models and Techniques for Its Study from Past to Present. In: Martin F., Stein T. and Howlin J. (eds) Mammary Gland Development. Humana Press, New York, NY, pp. 21–76. https://doi.org/10.1007/978-1-4939-6475-8_2

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. (2011) The American Journal of Human Genetics 88:76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.

Zhuang, Z., Ding, R., Peng, L., Wu J, Ye Y., *et al.* Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. BMC Genomics (2020) 21:344. https://doi.org/10.1186/s12864-020-6742-6