

## Article

# Genomic Prediction and Genome-Wide Association Study for Boar Taint Compounds

Sara Faggion <sup>1,\*</sup>, Elena Boschi <sup>1</sup>, Renata Veroneze <sup>2</sup>, Paolo Carnier <sup>1</sup> and Valentina Bonfatti <sup>1</sup>

<sup>1</sup> Department of Comparative Biomedicine and Food Science, University of Padova, Viale dell'Università 16, 35020 Padova, Italy; elena.boschi@phd.unipd.it (E.B.); paolo.carnier@unipd.it (P.C.); valentina.bonfatti@unipd.it (V.B.)

<sup>2</sup> Department of Animal Science, Universidade Federal de Viçosa, Viçosa 36570-999, Brazil; renata.veroneze@ufv.br

\* Correspondence: sara.faggion@unipd.it

**Simple Summary:** Some of the compounds involved in sexual steroids' metabolic pathways (i.e., androstenone, indole, and skatole) might accumulate in the adipose tissue of intact male pigs after sexual maturity, resulting in boar taint (BT). With a perspective future ban on surgical castration, in pig population where early slaughtering is not a viable option, the exploitation of genomic selection procedures might prevent this off-odor and off-flavor. The accuracy provided by the use of genomic information was equal or higher than the one obtained using pedigree information. This indicates that genomic selection could be beneficial for the traits investigated as it minimizes the need to collect individual measures of BT compounds. Several genomic regions, each with a small effect on BT compound concentrations, were identified. Genes previously associated with BT, reproduction traits, and fat metabolism are located in those genomic regions. Detection of candidate genes related to fat metabolism might be explained by the relationship between sexual steroid levels and fat deposition and be ascribed to the pig line investigated, selected for ham quality and not for lean growth.

**Abstract:** With a perspective future ban on surgical castration in Europe, selecting pigs with reduced ability to accumulate boar taint (BT) compounds (androstenone, indole, skatole) in their tissues seems a promising strategy. BT compound concentrations were quantified in the adipose tissue of 1075 boars genotyped at 29,844 SNPs. Traditional and SNP-based breeding values were estimated using pedigree-based BLUP (PBLUP) and genomic BLUP (GBLUP), respectively. Heritabilities for BT compounds were moderate (0.30–0.52). The accuracies of GBLUP and PBLUP were significantly different for androstenone (0.58 and 0.36, respectively), but comparable for indole and skatole (~0.43 and ~0.47, respectively). Several SNP windows, each explaining a small percentage of the variance of BT compound concentrations, were identified in a genome-wide association study (GWAS). A total of 18 candidate genes previously associated with BT (MX1), reproduction traits (TCF21, NME5, PTGFR, KCNQ1, UMODL1), and fat metabolism (CTSD, SYT8, TNNI2, CD81, EGR1, GIPC2, MIGA1, NEGR1, CCSER1, MTMR2, LPL, ERFE) were identified in the post-GWAS analysis. The large number of genes related to fat metabolism might be explained by the relationship between sexual steroid levels and fat deposition and be partially ascribed to the pig line investigated, which is selected for ham quality and not for lean growth.

**Keywords:** indole; skatole; androstenone; GWAS; genomic prediction; GBLUP; candidate genes

**Citation:** Faggion, S.; Boschi, E.; Veroneze, R.; Carnier, P.; Bonfatti, V. Genomic Prediction and Genome-Wide Association Study for Boar Taint Compounds. *Animals* **2023**, *13*, 2450. <https://doi.org/10.3390/ani13152450>

Academic Editor: Giuseppe Maiorano

Received: 5 June 2023

Revised: 13 July 2023

Accepted: 27 July 2023

Published: 28 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Boar taint (BT) is a “urine/faecal-like”, unpleasant flavor and odor released during cooking or heating of pork meat, and it occurs in male pigs after puberty due to the accumulation in the adipose tissue of lipophilic components such as androstenone, skatole, and indole [1]. The accumulation of those compounds depends on a number of factors,

including diet, age, sexual activity, and genetics [2–7]. Androstene (5 $\alpha$ -androst-16-en-3-one) is a steroid hormone produced by the testes, whereas skatole (3-methylindole) and indole (2,3-benzopyrrole) are produced by the bacterial breakdown of the amino acid tryptophane in the hind-gut [2–4,6].

Surgical castration of male piglets is normally performed within the first seven days of life of the animal and is largely used to prevent BT [8]. Also, castration avoids undesirable aggressive behaviors, generally observed in the entire male pigs, and meat and fat quality are better in castrates for dry-cured productions [9,10]. However, animal welfare concerns have led to the proposal of banning surgical castration in EU, driving farmers to grow intact males. Unfavorable consequences of surgical castration are also represented by (i) a marked reduction in animal feed efficiency, (ii) a lower lean meat content, and (iii) a higher content of saturated fatty acids, which, on one hand, makes castrates' fat firmer and reduces the process of rancidity during the maturation of dry cured products, but, on the other hand, reduces the nutritional value of fat [9,11,12].

Slaughtering intact male pigs before sexual maturity can prevent BT, but it is not a viable option for pigs intended for protected designation of origin (PDO) dry-cured ham production, since the minimum age at slaughter required by product specifications to meet quality standards is after sexual maturity [13,14].

Selective breeding might provide an effective and long-term approach to control BT, supporting the transition towards the gradual elimination of surgical castration in the pig production industry. According to Ducro-Steverink [15], selective breeding could decrease BT incidence from 30% to 10% in less than 5 years. Significant genetic variation has been reported for androstene, skatole, and indole, between and within pig lines. Heritability estimates are moderate to high, ranging from 0.46 to 0.72 for androstene, from 0.26 to 0.50 for skatole, and from 0.29 to 0.35 for indole [7,16–21]. Positive genetic correlations between the compounds responsible for BT suggest that selection for the reduction in one BT compound might be beneficial for the reduction in the others. Genetic correlations range from 0.19 to 0.61 between skatole and androstene [18,22–25], from 0.19 to 0.45 between indole and androstene, and from 0.51 to 0.62 between skatole and indole [18,22].

According to Lervik et al. [26], estimated breeding values (EBV) for androstene for different ages were reported to be positively correlated in Duroc boars aged 1–27 weeks. In addition, androstene is correlated with testosterone. These relationships indicate that EBV may be used as early indicators of the BT levels in adult pigs and that breeding to reduce androstene levels on the basis of EBV could have a significant impact on phenotype [26].

Routine BT compound phenotyping, required by traditional pedigree-based selection, is challenging and expensive, particularly when the number of selection candidates is large. In this context, the implementation of genomic selection procedures is greatly beneficial [27]. In some countries, BT has been successfully reduced through the use of genomic selection, and different breeding companies have produced “low-taint” breeding animals [20].

Genome-wide association studies (GWAS) revealed different quantitative trait loci (QTL) and candidate genes related to androstene, skatole, and indole. QTL affecting BT compounds have been found on most *Sus scrofa* chromosomes (1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 13, 14, and 15, as reported in Duarte et al. [7], and Botelho et al. [28]); this highlights the polygenic nature of BT, suggesting that genomic selection to decrease BT is the method of choice [29]. Detecting variants that significantly affect a given trait might improve genomic prediction accuracy of BT compounds [30], whereas the examination of post-GWAS data helps to identify candidate genes more effectively. When the GWAS detect multiple QTLs, gene network and pathway analyses contribute to elucidate the genetic regulation at the base of the biological processes underlying the traits and support the interpretation of GWAS results [31,32].

Several candidate genes, resulting from differential gene expression studies, were identified as encoding receptors and enzymes involved in the metabolism of androstenone, skatole, and indole in the liver and testis tissues [7]. Moreover, when comparing breeds and individuals with high and low levels of BT, discrepancies in the set of differentially expressed genes for androstenone, skatole, and indole levels were revealed [33]. Numerous studies showed that many of the differentially expressed genes in animals with different levels of BT compounds produced (1) phase I enzymes, such as the cytochrome P450 (Cyt-P450) family and their isoforms, (2) phase II enzymes, such as the sulfotransferases for steroid conjugation, and (3) enzymes in the testes that are responsible for the production of androstenone substrates and steroids [34–37].

So far, studies have focused on pig lines selected for lean growth, but since there is a well-known relationship between steroid hormone levels and adipogenesis [38], it is of interest to investigate associations between BT and genomic regions in pig populations where the breeding goals are addressed to enhanced ham quality and includes, with positive weights, traits related to adipogenesis.

Hence, the aims of the study were: (1) to evaluate the accuracy of traditional pedigree-based and genomic-enabled methods in the prediction of BT compound concentrations in a purebred sire line selected for carcass and ham quality traits; (2) to identify QTL regions and candidate genes associated with BT compound concentrations, through a single-trait GWAS approach.

## 2. Materials and Methods

### 2.1. Animals and Sample Collection

The study was conducted on 1075 intact young boars of the C21 Goland sire line (Gorzagri, Fonzaso, Italy), selected to enhance carcass and ham quality traits and widely used in Italy to generate commercial heavy pigs intended for PDO dry-cured ham production. For BT compounds quantification, a subcutaneous adipose tissue sample of approximately 0.5 g was collected from the neck area of each young boar (subjected to local anesthesia) using a biopsy device (SUISAG, Sempach, Switzerland [39]); samples were kept at  $-80^{\circ}\text{C}$  until the laboratory analysis. Samples were collected between January 2020 and June 2022. The average animal age and body weight at biopsy sampling were  $209 \pm 12$  d and  $153.48 \pm 15.46$  kg, respectively. Boar taint compound concentrations in adipose tissue is highly representative of the concentration of those compounds in plasma. Several studies have underlined how androstenone and skatole are easily transferred from plasma to the adipose tissue, such that if the concentration of those compounds in peripheral plasma increases, a high accumulation of androstenone and skatole in the adipose tissue is expected [40–42].

### 2.2. Boar Taint Compound Quantification

Reversed-phase HPLC with fluorescence detection was used to quantify BT compound concentrations. Sample preparation was performed following the method by Rostellato et al. [18]. Two different chromatographic separations were performed as described by Boschi et al. [43]: (1) an isocratic separation for skatole (SKA) and indole (IND) and (2) a gradient separation for androstenone (AND). The total time of the isocratic separation for SKA and IND quantification was 5 min. For AND quantification, the sample was injected after sample derivatization following the method described by Rostellato et al. [18]. The total time of the chromatographic separation was 11 min. As reported by Boschi et al. [43], the detection limit of the analysis was  $0.08\text{ }\mu\text{g/g}$ ,  $1.21\text{ ng/g}$ , and  $1.34\text{ ng/g}$  for androstenone, skatole, and indole, respectively. The residual coefficient of variation (CV) of compound concentration across days was  $<5.2\%$ , whereas the residual CV of compound concentration within day of analysis was  $<2.6\%$ .

The frequency distribution of BT compound concentrations, measured as ng/g for indolic compounds and as µg/g for androstenedione, were skewed and then normalized through log-transformation of the original data.

### 2.3. Genotyping

A saliva sample was collected from each of the 1075 pigs and sent to GeneSeek Inc. laboratory (Lincoln, NE, USA) for DNA extraction and genotyping. Animals were genotyped using a high-density SNP chip (GGP Porcine 50K). All the animals were successfully genotyped. SNP markers with a minor allele frequency (MAF) lower than 1% and call rate lower than 95% were discarded. Missing genotypes were then imputed using the FImpute software [44] exploiting genotype data and pedigree information available for the C21 line. The final number of available SNP genotypes per animal was 29,844.

### 2.4. Genomic Predictions

Breeding values (EBV) for BT compound concentrations were predicted with pedigree-based BLUP (PBLUP) and genomic BLUP (GBLUP) methods. Variance components, heritability and pedigree-based EBVs were estimated using the Average Information Restricted Maximum Likelihood method implemented in the BLUPf90+ suite of programs [45] with the following univariate mixed linear model:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\mathbf{b} + \mathbf{W}\mathbf{a} + \mathbf{e} \quad (1)$$

where  $\mathbf{y}$  is the vector of phenotypes (i.e., the log-transformed concentration of one BT compound in adipose tissue);  $\mu$  is the average;  $\mathbf{1}$  is an all-ones vector;  $\mathbf{b}$  is the random effect of day of BT compound analysis (91 levels);  $\mathbf{X}$  is an incidence matrix relating  $\mathbf{y}$  to  $\mathbf{b}$ ;  $\mathbf{a}$  is a vector of random animal additive genetic effects assumed to be  $N(0, A\sigma_a^2)$ , where  $N(\cdot)$  indicates a Gaussian probability density function,  $A$  is the numerator relationship matrix between animals (from pedigree data) and  $\sigma_a^2$  is the additive genetic variance;  $\mathbf{W}$  is an incidence matrix relating  $\mathbf{y}$  to  $\mathbf{a}$ ; and  $\mathbf{e}$  is a vector of random residuals assumed to be  $N(0, I\sigma_e^2)$ , where  $I$  is an identity matrix of appropriate order and  $\sigma_e^2$  is the residual variance. The same model was used to estimate the genomic-based EBVs (GEBVs) by GBLUP replacing the  $A$  relationship matrix with the genomic relationship matrix ( $G$  matrix) [46]. The day of BT compound analysis has been included in the model because it allowed to take into account the small variations in sample derivatization due to different reagent batches, handling and preparation.

Accuracy of PBLUP and GBLUP models was assessed in a 5-fold cross-validation. Firstly, phenotypes (log-transformed AND, IND, and SKA concentration) were adjusted for the random effect of day of BT compound analysis ( $\mathbf{y}_{adj}$ ) using PREDICTF90 from the BLUPf90+ suite of programs [45] using model (1). The set of data was then randomly split into five equally-sized subsets. At each round, 4 out of the 5 subsets were used as a training set, masking the phenotypes of the subset that was used as a validation set, and EBVs and GEBVs were predicted for the individuals with masked phenotype. This procedure was repeated until each of the five subsets were used as a validation set. The accuracy ( $R_{EBV,BV}$ ) of PBLUP and GBLUP predictions was estimated as follows:

$$R_{EBV,BV} = \frac{r_{EBV,y_{adj}}}{\sqrt{h^2}} \quad (2)$$

where  $r_{EBV,y_{adj}}$  is the Pearson correlation between the predicted EBVs and the adjusted phenotypes ( $\mathbf{y}_{adj}$ ) and  $h^2$  is the trait heritability. Prediction bias was assessed as the regression coefficient (slope,  $\beta$ ) of the EBVs on the adjusted phenotypes. The Hotelling–Williams test [47] was used to test the difference in accuracy between PBLUP and GBLUP predictions.

### 2.5. Genome-Wide Association Studies for Boar Taint Compounds

A genome-wide association study (GWAS) was performed to test the association between the phenotype of each BT compound and SNP genotypes. The analyses were performed following a GBLUP GWAS procedure [48,49] using the BLUPf90+ suite of programs [45]. First, GEBV were predicted using BLUPf90+ using the variance components estimated with the pedigree-based BLUP. The analysis was performed according to the single-trait animal model (1), but the vector of random animal additive genetic effects was assumed to be distributed as  $N(0, G\sigma_a^2)$ , with  $N(\cdot)$  indicating a normal probability density function,  $G$  being the genomic relationship matrix, and  $\sigma_a^2$  being the additive genetic variance.

$G$  was obtained as in VanRaden [46]:

$$G = \frac{ZZ'}{2 \sum_{i=1}^m p_i(1 - p_i)} \quad (3)$$

where  $Z$  is a matrix of gene content (0, 1, 2) adjusted for allele frequencies;  $m$  is the number of SNPs, and  $p_i$  is the minor allele frequency of  $i$ -th SNP.

The allele substitution effect of each SNP was then derived from GEBV using post-GSf90 [50]:

$$\hat{g} = \frac{1}{2 \sum_{i=1}^m p_i(1 - p_i)} Z' G^{-1} \hat{u} \quad (4)$$

where  $\hat{g}$  is the vector of SNP effect and  $\hat{u}$  is the vector of GEBV.  $p$ -values were obtained as described in Aguilar et al. [51]:

$$p - \text{value}_i = 2 (1 - \Phi \left( \left| \frac{\hat{g}_i}{\text{sd}(\hat{g}_i)} \right| \right)) \quad (5)$$

where  $\Phi$  is the cumulative standard normal function and individual prediction error variances of SNP effect estimates were obtained as in Gualdrón-Duarte et al. [48]. In order to address the issue of multiple testing, the genome-wide significance threshold was adjusted at  $p\text{-value} = 0.05/m$ , thus applying a stringent Bonferroni correction.

Explained percentage of genetic variance was calculated using subsequent windows of adjacent SNPs within a distance of 0.4 Mb, which is the average haplotype block size in commercial pig lines [52]. Manhattan plots of GWAS results and plots of the variance explained by the SNP windows were obtained with the R package qqman [53].

### 2.6. Identification of QTLs and Candidate Genes

For each BT compound, the 0.4 Mb windows that explained 0.5% or more of the total genetic variance were selected were considered as putative QTL and selected for candidate genes search. The threshold of 0.5% was chosen based on the literature [54,55] and on the expected contribution of SNP windows [56]. Assuming that the totality of the windows explained 100% of the genetic variance and assuming an equal contribution of all windows, the expected proportion of genetic variance explained by each window would be equal to  $100/N$ , where  $N$  is the total number of windows obtained for each trait. In total, 4040, 4037, and 4050 windows were obtained for AND, IND, and SKA, respectively; hence, the expected proportion of genetic variance explained by each window, given the assumptions stated above, would be equal to 0.025% for each trait. The threshold of 0.5% used in the present study is equal to 20 times the expected variance explained ( $0.025\% \times 20 = 0.5\%$ ).

The Ensembl VEP (variant effect predictor) tool (<https://www.ensembl.org/info/docs/tools/vep/index.html>; accessed on 16 September 2022) was used to determine the effect of SNPs on genes, transcripts, protein sequence and regulatory regions. Genes located within 200 kb upstream and downstream of the lead SNP were considered putative candidate genes associated with the trait. The candidate genes nearby the genome-wide significant SNPs were identified by Ensembl database (<https://www.ensembl.org/>; accessed on 16 September 2022) using the gene annotation information of *Sus*

*scrofa* reference genome (Sscrofa11.1), which is publicly available at NCBI. The genomic locations were downloaded from PigQTL database (<https://www.animalgenome.org/pig/>; accessed on 20 September 2022) and genes within the candidate region were considered as candidate genes.

### 3. Results and Discussion

#### 3.1. Descriptive Statistics for Androstenone, Indole and Skatole Concentrations

For each BT compound, the number of animals with phenotypic data, and the sample mean and standard deviation before and after log-transformation are reported in Table 1. The threshold of the compound concentrations above which carcasses are generally defined as tainted are 1.0 µg/g for AND [57,58] and 0.20–0.25 µg/g for SKA [9]; in our study, only 58.25% of the data were below the androstenone threshold, whereas 96.25% of the data were below the skatole threshold. Narrow sense heritabilities (Table 1) for BT compounds estimated from pedigree-based relationships on our experimental population were moderate ( $h^2 \pm SE$ ):  $0.30 \pm 0.08$  for AND,  $0.43 \pm 0.09$  for IND,  $0.51 \pm 0.09$  for SKA. The estimates indicate that moderate to high genetic gain could be obtained through selection against boar taint compounds in the evaluated population. These estimates are comparable to those reported in the literature: in Landrace, Large White, Duroc, and Pietrain breeds, estimates of the genetic parameters ranged between 0.39 and 0.73 for AND and between 0.32 and 0.69 for SKA [20]; in the pig line used in the current study, heritability estimates were 0.39 for AND and 0.60 for SKA [18]. Even though estimates for IND are scarce in the literature, overall, they are comparable ( $0.33$  [17];  $0.39$  [18]) or slightly higher ( $0.55$  [39]) than the one obtained in the current study.

**Table 1.** Number of animals with phenotypic data, mean ( $\pm$ SD) concentration of boar taint compounds before and after log-transformation, and heritability ( $h^2 \pm SE$ ) of the investigated traits.

Trait	N	Mean ( $\pm$ SD) before Log-Transformation	Mean ( $\pm$ SD) after Log-Transformation	$h^2$
Androstenone (ng/g)	1051	1162.8 $\pm$ 1018.2	6.77 $\pm$ 0.76	0.30 $\pm$ 0.08
Indole (ng/g)	1073	21.7 $\pm$ 25.0	2.63 $\pm$ 0.93	0.43 $\pm$ 0.09
Skatole (ng/g)	1069	48.3 $\pm$ 70.6	3.31 $\pm$ 1.04	0.51 $\pm$ 0.09

#### 3.2. Genomic Predictions

The Pearson correlations between predicted EBV and adjusted phenotypes ( $r_{EBV,y_{adj}}$ ) ranged between 0.20 (AND) and 0.33 (SKA) for pedigree-based BLUP and between 0.27 (IND) and 0.35 (SKA) for genomic-based BLUP (Table 2). The slope of the regression of  $y_{adj}$  on EBV ( $\beta$ ), ranged between 0.89 (AND) and 1.00 (SKA), whereas GEBV tended to be inflated ( $\beta < 0.89$ ). The accuracies ( $R_{EBV,BV}$ ) ranged between 0.36 (AND) and 0.45 (IND) for pedigree-based BLUP and between 0.42 (IND) and 0.58 (AND) for genomic-based BLUP (Table 2).

**Table 2.** Pearson correlations between predicted EBV and adjusted phenotypes ( $r_{EBV,y_{adj}}$ ), corresponding accuracies ( $R_{EBV,BV}$ ) and  $\beta$  coefficients.

Trait	$r_{EBV,y_{adj}}^1$		$R_{EBV,BV}^2$		$\beta$	
	PBLUP	GBLUP	PBLUP	GBLUP	PBLUP	GBLUP
Androstenone	0.198	0.319 **	0.362	0.582 **	0.893	0.887
Indole	0.292	0.273	0.446	0.417	1.008	0.663
Skatole	0.326	0.354	0.452	0.491	0.963	0.768

<sup>1</sup> Pearson correlation between predicted EBV and adjusted phenotypes ( $y_{adj}$ ). <sup>2</sup> Accuracy computed as in Equation (2); \*\* indicates statistically significant differences ( $p < 0.01$ ) between PBLUP and GBLUP.

For AND, GBLUP performed significantly better than PBLUP, whereas for IND and SKA, the accuracies of PBLUP and GBLUP were not significantly different. The accuracies obtained from GBLUP for AND and SKA are in agreement with literature estimates. De Campos et al. [19] reported accuracies, measured as the correlation between the observed and the predicted phenotype divided by the square root of trait heritability, equal to 0.63 and 0.57 for AND and SKA, respectively, using a Ridge Regression BLUP method. Lukić et al. [59] and Botelho et al. [60] estimated the genomic breeding values for AND and SKA through a GBLUP approach with correlations between the predicted breeding values and the phenotypes of 0.27–0.35 and 0.21–0.49 for AND and SKA, respectively. In that study, the accuracies adjusted for the square root of heritability (0.56 and 0.76 for AND and SKA, respectively) were much higher than those obtained in this study.

In the literature, only Botelho et al. [60] have reported results of genomic prediction for IND: for a dataset consisting of records from a single-line pig population, prediction accuracy computed as the Pearson's correlation between GEBV and phenotypes adjusted for fixed effects ranged between 0.24 and 0.26, whereas for a multi-line pig population, correlations were slightly lower (0.21–0.26). Those correlations are similar to that obtained in the present study.

### 3.3. Single-Trait Genome-Wide Association Study for Boar Taint Compounds

The genome-wide significance threshold after Bonferroni correction and after  $-\log_{10}$  transformation was estimated to be 5.78. The genome-wide association study failed to identify any genome-wide QTL associated to AND, IND, or SKA concentrations (Figure 1). In previous studies performed on single-line pig populations, several markers associated with variation in BT compound concentrations were detected: 33 genomic regions associated with the variation in at least one of the BT components were reported by Große-Brinkhaus et al. [29]; 14 regions for AND and 10 for SKA, in Landrace, and 14 regions for AND and 4 for SKA, in Duroc, were detected by Grindflek et al. [22].

The ability of a GWAS to identify genomic regions significantly associated with a trait variation depends on the power of the experiment, which is related to sample size and phenotype variability. The number of samples included in our study is considered sufficient for a GWAS, but might not be large enough to ensure the identification of DNA regions with a small effect on the trait. In addition, phenotype variation in our dataset was remarkably lower than that reported in other studies [22,29]. Both these aspects might have impacted our ability to detect significant associations.

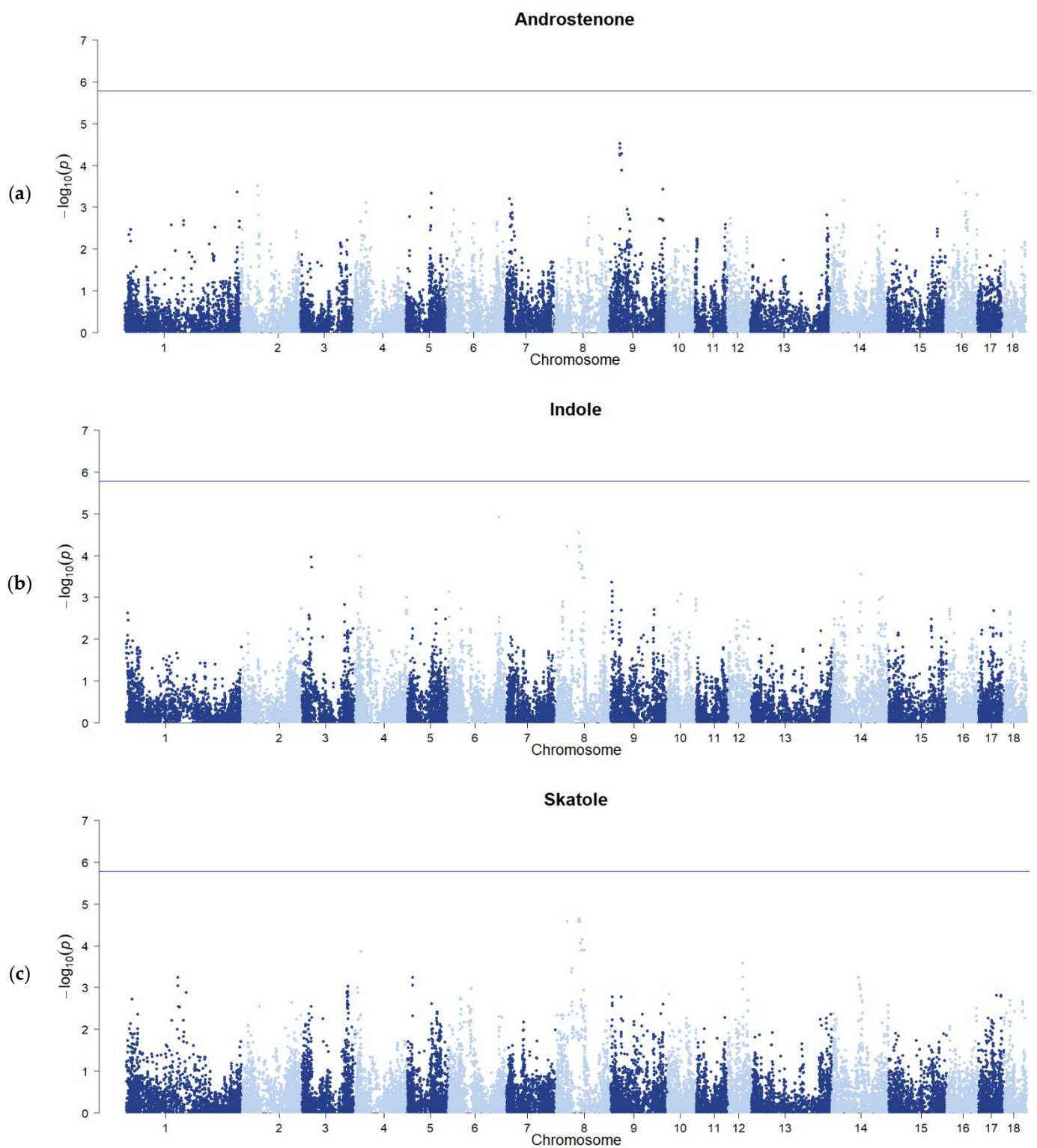
When hypothesis tests are performed using a high number of SNPs, determining the multiple testing correction threshold in GWAS might be challenging; this is typically performed using the Bonferroni correction. Bonferroni correction is a highly conservative method, particularly when datasets consist of a high number of SNPs and the independence assumption does not hold, which results in a high penalty for GWAS [61]. To test the associations between BT compounds and SNP markers, different studies [28,60] have proposed an approach based on the percentage of variance explained to select markers potentially associated to BT compounds. The threshold of 0.5% was chosen based on the literature [54,55] and on the expected contribution of SNP windows [56]. Assuming an equal contribution of all windows, the expected proportion of genetic variance explained by each window would be equal to 0.025% for all the traits ( $100/N$ , where  $N$  is the total number of windows; we obtained 4040, 4037, and 4050 windows for AND, IND, and SKA, respectively). The threshold of 0.5% used in the present study is equal to 20 times the expected variance explained ( $0.025\% \times 20 = 0.5\%$ ).

Following this strategy, we have detected signals from several chromosome regions, most of them consistent with the literature, as well as novel chromosome regions. The percentage of variance explained by each SNP windows for AND, IND, and SKA is depicted in Figure 2. For all BT compounds, multiple genomic windows were detected, each window explaining a small amount of the trait additive genetic variance, which supports

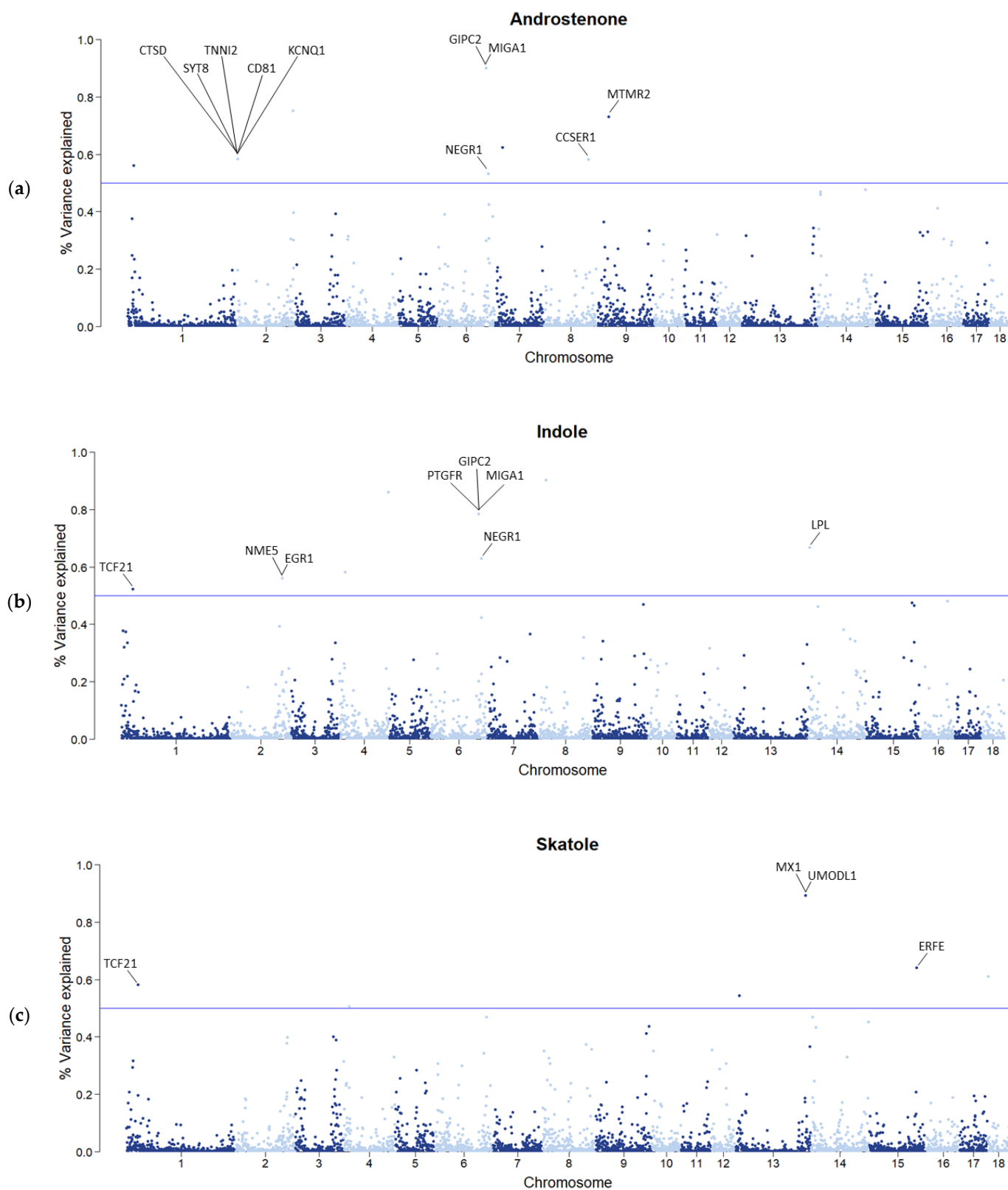
the hypothesis of polygenic nature of these traits [28,62–64]. In particular, 18 windows explaining more than 0.5% of the genetic variance were detected.

In a previous study performed on a single-line pig population, 37 SNPs explaining 13.7% of the variance for AND were reported [65]. Studies focused on pigs from more than one sire line led to detection of a lower number of SNP windows (16), explaining in total a small amount of genetic variance (2.90%, 2.15%, and 4.47% for AND, SKA, and IND, respectively [28]). Several authors hypothesized that variations in population stratification, genetic backgrounds, linkage disequilibrium, population size, and other factors could explain differences across studies [50,55,66]. Moreover, variation in BT varies across breeds, probably due to different selection goals for each breed. Also, a different expression of candidate genes linked to AND, SKA, and IND has been reported for individuals characterized by different levels of BT compounds or belonging to various breeds [33].





**Figure 1.** Genome-wide association plots for androstenone (a), indole (b), and skatole (c) concentrations. The horizontal line identifies the genome-wide significance threshold ( $-\log_{10}(0.05/m)$ , with  $m$  as the total number of markers).



**Figure 2.** Percentage of genetic variance explained by SNP windows for androstenone (a), indole (b), and skatole (c) concentrations. Each dot represents one SNP window of 0.4 Mb. The horizontal line identifies the threshold of 0.5% used to select SNPs for candidate gene search.

### 3.4. Identification of QTLs

For AND, QTL regions exceeding the 0.5% threshold of explained variance were identified on *Sus scrofa* chromosome 1, 2, 6, 7, 8, and 9 (Figure 2a, Table 3). Windows explained between 0.53 and 0.90% of the variance; the most relevant window was located on SSC6. Together, the eight windows identified for AND explained 5.25% of the genetic variance of the trait. Several studies have reported QTL regions associated with AND on SSC1 [28,65], 2 [28,63], 6 [22,28,62,63,65,67,68], 7 [62,63,69], and 9 [62,63], whereas the one on SSC8 has never been reported before.

**Table 3.** QTL regions identified for androstenone, indole, and skatole. *Sus scrofa* chromosome (SSC), position of QTL region, number of SNPs within the QTL region and percentage of variance explained by the windows are reported.

Trait	SSC	QTL Region (Mb)	N SNP	Variance Explained (%)	Other Associated Boar Taint Compound
Androstenone	1	15.05–15.44	14	0.56	
	2	1.23–1.54	13	0.58	
	2	156.30–156.65	12	0.75	
	6	135.01–135.40	15	0.90	Indole
	6	140.66–141.06	12	0.53	Indole
	7	18.15–18.54	11	0.62	
	8	128.63–128.99	16	0.58	
	9	27.88–28.27	12	0.73	
Indole	1	29.76–30.15	9	0.52	Skatole
	2	140.25–140.64	16	0.56	
	4	15.44–15.74	8	0.58	Skatole
	4	141.08–141.46	9	0.86	
	6	134.85–135.24	15	0.78	Androstenone
	6	140.95–141.31	14	0.63	Androstenone
	8	17.93–18.32	11	0.90	
	14	3.77–4.13	14	0.67	
Skatole	1	30.00–30.38	13	0.58	Indole
	4	15.44–15.74	8	0.50	Indole
	13	9.55–9.94	23	0.54	
	13	204.99–205.38	16	0.89	
	15	137.53–137.93	10	0.64	
	18	0.76–1.12	9	0.61	

For IND, windows located on SSC1, 2, 4, 6, 8, and 14 explained between 0.52 and 0.90% of the variance, with the most relevant windows located on SSC4 and 8 (Figure 2b, Table 3). The eight windows identified explained 5.50% of the genetic variance of IND. Besides the QTL regions that have been previously detected as associated with IND or pork odor (SSC2 and 6 [28,70]; SSC14 [63]), we detected additional QTLs on SSC1 and 8.

Six windows explaining 3.76% of the variance of SKA were identified on SSC1, 4, 13, 15, and 18 (Figure 2c, Table 3). Each window explained between 0.50% (on SSC4) and 0.89% (SSC13) of the variance. For SKA, one chromosome reported in the present study (SSC4) has been previously associated with the intensity of smell and taste of pork [67], whereas the QTL region in SSC13 has been associated with SKA [28]. In the current study, additional QTLs associated with SKA have been detected on SSC1, 15, and 18.

One window, common to AND and IND, located on SSC6 explained 0.53% and 0.63% of the genetic variance of the two traits, respectively. Two windows on SSC1 and SSC4 were common to IND and SKA; those windows explained 0.52% and 0.58% (SSC1) and 0.58% and 0.50% (SSC4) of the genetic variance, for IND and SKA, respectively.

### 3.5. Candidate Genes

Considering the genomic regions associated with either AND, IND, or SKA (Table 4), we identified 18 candidate genes encoding receptors and enzymes associated with BT, reproduction traits, and fat metabolism. Only one of the genes identified (MX1, *Myxovirus resistance 1*, involved in the resistance to influenza virus in pigs [71]) has been previously reported in the literature in association with BT, in particular, with the concentration of AND [37]. In the current study, the gene explained 0.89% of the variance of SKA.

The large majority of the GWAS performed on BT [22,28,29,65,72,73] identified associations with the genes encoding Cytochrome P450 family enzymes (CYP) that regulate steroid metabolic processes and are known to play a key role in the metabolism of BT compounds in the liver [5]. Other enzymes with a well-known role in BT compound metabolism (e.g., SULT [5]) have been identified by GWAS by Duijvesteijn et al. [65] and Gregersen et al. [72]. Only a few studies investigated the variance explained by SNP windows [28,29]. However, estimates of the variance explained obtained with different methods are not comparable. In the study by Botelho et al. [28], who estimated the variance explained by SNP windows using the same approach as the one used in the current study, the variance explained by the CYP candidate gene CYP27B1 was 0.56% for IND, and 0.01% for AND and SKA. This indicates that even the genes directly involved in AND, IND, and SKA metabolism have a small contribution for the trait phenotypic variance.

In our study, several genes associated with sexual development (TCF21, NME5, PTGFR, KCNQ1, UMODL1) were identified. TCF21 (*Transcription Factor 21*) is involved in testis growth and development in chicken [74]. In mice, TCF21 is also the first direct downstream target of the male *sex determining factor* (SRY) and both have similar effects on Sertoli cell differentiation and embryonic testis development in rats [75]. NME5 (*Nucleoside diphosphate kinase homolog 5*, also known as *NME/NM23 family member 5 gene*) is considered the best candidate for sperm concentration in Landrace [55]. According to Munier et al. [76], this gene is highly and specifically expressed in testis and the encoded protein is important for the initial stages of spermatogenesis. Choi et al. [77] reported that, when expression of NME5 is reduced, the spermatids in the testes become more sensitive to oxidative stress, leading to DNA damage and reduced sperm cell numbers, showing that this gene plays a crucial role in spermiogenesis. PTGFR (*Prostaglandin F Receptor*) is known to affect sexual behavior and to stimulate testicular hormone secretion in boars [78,79]. Epimutations (hypermethylation) in the KCNQ1 (*Potassium voltage-gated channel subfamily Q member 1*) gene have been reported in association with poor semen traits or male infertility [80,81].

One of the identified genes (UMODL1, *Uromodulin Like 1*) has been reported to be associated with female fertility, and, in particular, with oocyte fecundity. It is a gonadotropin-responsive gene expressed in both oocytes and the thymic medulla in mice. Mice carrying an extra allele of UMODL1 have drastically reduced litter sizes and increased premature infertility [82]. Reports suggest that disruptions to UMODL1 function may accelerate ovarian senescence [82]. As AND synthesis is stimulated along with other steroids, a relationship between BT and reproduction traits has been hypothesized. Genetic correlations between AND concentration and several female reproduction traits, including the total number of piglets born, have been reported in the literature [7].

Several candidate genes (CTSD, SYT8, TNNT2, CD81, EGR1, GIPC2, MIGA1, NEGR1, CCSER1, MTMR2, LPL, ERFE) related to adipocyte biogenesis, lipid metabolic processes and deposition, and to variations in fat thickness and intramuscular fat were also identified. CTSD (*Cathepsin D*) encodes for an aspartic lysosomal proteinase. Transcription of this gene is initiated from several sites, including one which is a start site for an estrogen-regulated transcript. The gene has been associated with growth, fat deposition, and production traits in the Italian Large White pig population [83,84]. It is worth mentioning that high cathepsin activity of porcine skeletal muscle is correlated to defects of dry-cured hams associated with excessive meat softness [85] and tainted hams have been associated to increased proteolysis and impaired organoleptic quality [86]. *Synaptotagmin 8* (SYT8)

encodes a member of the synaptotagmin protein family, which includes membrane proteins involved in neurotransmission and hormone secretion. The  $\text{Ca}^{2+}$  non-binding isoform syt 8 has been detected in insulin-secreting cells. Glucose stimulated expression of SYT8 in human islets and SYT8 knock-down have shown to impair insulin release [87], even though the role of syt 8 in insulin secretion remains unclear [88].

TNNI2 (*Troponin I2*) gene has been associated with fat percentage, lean meat percentage, loin eye area, thorax–waist backfat thickness, and average backfat thickness in pigs [89]; TNNI2 gene affected significantly muscle pH, color, marbling, and intramuscular fat content [90]. CD81 (*Cluster of differentiation 81*) is a beige adipocyte progenitor cell marker and is required for de novo beige fat biogenesis; “beiging” of white adipose tissue is an adaptive process initiated as a response to cold temperatures in which numerous mitochondria-enriched thermogenic adipocytes with multi-locular lipid droplets (i.e., beige adipocytes) emerge within white adipose tissue [91]. EGR1 (*Early growth response 1*) influences adipocyte differentiation and was found to have an anti-adipogenic action [92], acting in opposition to EGR2. The regulation of both is required for maintaining appropriate levels of adipogenesis. *Neuronal growth regulator 1* (NEGR1), *Mitoguardin 1* (MIGA1, also known as protein FAM73A), and *GIPC PDZ domain containing family member 2* (GIPC2), are responsible for genetic predisposition to common forms of obesity in human, especially those resulting from deposition of subcutaneous fat, and they have been associated to subcutaneous fat thickness in a combined human and pig study [93]. *Coiled-Coil Serine Rich Protein 1* (CCSER1) was identified as a candidate gene for backfat thickness in Chuying-black pigs [94]. MTMR2 (*Myotubularin related protein 2*) was proposed as a functional candidate gene for intramuscular fat content in GWAS and signatures of selection studies in a Duroc pig population selected for intramuscular fat content [95]. *Lipoprotein lipase* (LPL) can hydrolyze the lipoprotein triacylglycerols from cycling and release free fatty acid for skeletal muscle uptake. The increased expression of the gene indicates that the ability to uptake exogenous fatty acid and synthesize triacylglycerols could be increased, which could intensify fat deposition [96]. ERFE (*Myonectin*, also known as *Erythroferrone*) is associated with iron metabolism, but is also a key player in regulating local and systemic lipid metabolism [97]. In particular, when fed a high-fat diet, myonectin-knockout mice had significantly elevated very-low-density-lipoprotein–triglyceride and impaired lipid clearance from circulation following an oral lipid load. Fat distribution between adipose tissue and liver was also altered. Greater fat storage in adipocytes was associated with increased postprandial lipoprotein lipase activity in adipose tissue [97].

The large number of candidate genes related to fat metabolism can be ascribed to the well-known relationship between sexual steroid levels and fat deposition. According to Kelly and Jones [35], reduced tissue testosterone facilitates triglyceride storage in adipocytes by allowing increased lipoprotein lipase activity and stimulating pluripotent stem cells to mature into adipocytes. In addition, increased adipocyte mass is associated with greater insulin resistance and further inhibits testosterone synthesis [38]. Due to the relationship between BT compound concentrations and fat deposition, the pig line investigated, which is selected for ham quality and not for lean growth, might have influenced the results obtained and might explain the discrepancies in the identified genes between our study and the literature. In a recent study performed on the same pig line, significant genetic correlations were detected between BT compound concentrations measured in purebred pigs and carcass backfat or ham subcutaneous fat measured in crossbred heavy pigs [98].

**Table 4.** Candidate genes identified for androstenone (AND), indole (IND), and skatole (SKA), considering the genomic regions associated; *Sus scrofa* chromosome (SSC), position of the gene, and gene name and functions are reported.

Trait	SSC	Position (Mb)	Gene	Function	References
IND, SKA	1	29.98–29.98	TCF21	Associated with testis growth and development	[74]
AND	2	1.19–1.20	CTSD	Encodes lysosomal aspartyl proteinase	[83]
AND	2	1.25–1.25	SYT8	Encodes a member of the synaptotagmin protein family (neuro-transmission and hormone secretion)	[87,88]
AND	2	1.25–1.25	TNNI2	Associated with intramuscular fat content	[89,90]
AND	2	1.63–1.64	CD81	Associated with adipocyte biogenesis	[91]
AND	2	1.68–1.99	KCNQ1	Associated with male infertility	[80]
IND	2	140.13–140.17	NME5	Associated with spermiogenesis	[76]
IND	2	140.44–140.45	EGR1	Associated with adipocyte differentiation	[92]
IND	6	134.73–134.77	PTGFR	Binds PGF2 $\alpha$ , associated with prostaglandin F receptor, affects sexual behavior, stimulates testicular hormone secretion	[78,79]
AND, IND	6	135.08–135.22	GIPC2	Associated with subcutaneous fat thickness	[93]
AND, IND	6	135.42–135.51	MIGA1	Associated with backfat thickness	[93]
AND, IND	6	140.78–141.65	NEGR1	Associated with subcutaneous fat thickness	[93]
AND	8	127.73–128.95	CCSER1	Associated with backfat thickness and loin depth	[94]
AND	9	27.85–27.97	MTMR2	Associated with lipid metabolic processes	[95]
IND	14	4.11–4.14	LPL	Hydrolyses triacylglycerols in plasma lipoproteins	[96]
SKA	13	204.84–204.87	MX1	Associated with boar taint	[37]
SKA	13	205.43–205.49	UMODL1	Involved in the regulation of ovarian follicle development	[82]
SKA	15	137.73–137.74	ERFE	Promotes lipid uptake into adipocytes and hepatocytes	[97]

#### 4. Conclusions

In the current study, the accuracy in predicting EBV using only pedigree data or only genomic data was compared. For androstenone, the exploitation of genomic data led to significantly more accurate EBV predictions compared to those obtained using pedigree data, whereas for IND and SKA, the accuracies were not significantly different. In general, genomic estimated breeding values can be used in selective breeding aimed to decrease boar taint compounds in entire male pigs, avoiding the challenging and expensive routine BT compound phenotyping required by traditional pedigree-based selection. Hence, genomic selection represents a promising alternative to male piglet castration and allows pigs to be slaughtered at the age and body weight required by product specifications of protected designation of origin dry-cured hams. Genome-wide association studies revealed that boar taint compounds are influenced by a large number of loci, each explaining a small percentage of the variance. The polygenic nature of boar taint makes genomic selection strategies particularly beneficial as genomic BLUP ignores any information about the genetic architecture of a trait and assumes that each marker equally contributes to the genetic variance. Post-GWAS analyses of biological processes led to the identification of 18 candidate genes encoding receptors and enzymes associated with BT (MX1), reproduction traits (TCF21, NME5, PTGFR, KCNQ1, UMODL1), and fat metabolism (CTSD, SYT8, TNNI2, CD81, EGR1, GIPC2, MIGA1, NEGR1, CCSER1, MTMR2, LPL, ERFE). The large number of candidate genes (12 out of 18) related to fat metabolism might be explained by the well-known relationship between sexual steroid levels and fat deposition and be, in part, ascribed to the pig line investigated, which is selected for ham quality and not for lean growth.

**Author Contributions:** Conceptualization, V.B. and P.C.; methodology, S.F., E.B. and R.V.; software, S.F.; validation, S.F.; formal analysis, S.F.; investigation, E.B., S.F. and R.V.; resources, E.B.; data curation, S.F. and V.B.; writing—original draft preparation, S.F.; writing—review and editing, S.F., V.B. and P.C.; visualization, S.F.; supervision, V.B., P.C. and R.V.; project administration, V.B.; funding acquisition, V.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study received funds from the project “Applied phenomics and genomics in pigs for the identification and use of new phenotypes in breeding plans—PigPhenomics”, funded by the Italian Ministry of Education, University and Research (MIUR), within the PRIN 2017 research programme (Research Projects of National Relevance) grant n. 201759JRPY.

**Institutional Review Board Statement:** Ethical review and approval were obtained by the Local Health and Social Care Service n. 8 (Servizio Veterinario di Igiene degli Allevamenti e delle Produzioni Zootecniche, Montebelluna, Treviso, Italy) after transmission of procedures for in vivo collection of adipose tissue samples through biopsies (Protocollo 26B/2012). To guarantee animal welfare and health, a trained veterinarian performed all sampling procedures using local anesthesia.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Restrictions apply to the availability of these data. Data were obtained from Gorzagri (Fonzaso, Italy) and are available from the authors with the permission of Gorzagri.

**Acknowledgments:** The authors wish to thank Gorzagri (Fonzaso, Italy) for having provided the animals used in the study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Wesoly, R.; Weiler, U. Nutritional influences on skatole formation and skatole metabolism in the pig. *Animals* **2012**, *2*, 221–242.
2. Claus, R.; Weiler, U.; Herzog, A. Physiological aspects of Androstenone and skatole formation in the Boar—A review with experimental data. *Meat Sci.* **1994**, *38*, 289–305.
3. Babol, J.; Squires, E.J.; Gullett, E.A. Factors affecting the level of boar taint in entire male pigs as assessed by consumer sensory panel. *Meat Sci.* **2002**, *61*, 33–40.
4. Aldal, I.; Andresen, Ø.; Egeli, A.K.; Haugen, J.-E.; Grødum, A.; Fjetland, O.; Eikaas, J.L. Levels of androstenone and skatole and the occurrence of boar taint in fat from young boars. *Livest. Prod. Sci.* **2005**, *95*, 121–129.
5. Zamaratskaia, G.; Squires, E.J. Biochemical, nutritional and genetic effects on boar taint in entire male pigs. *Animal* **2009**, *3*, 1508–1521.
6. Aluwé, M.; Millet, S.; Bekaert, K.M.; Tuytens, F.A.M.; Vanhaecke, L.; De Smet, S.; De Brabander, D.L. Influence of breed and slaughter weight on boar taint prevalence in entire male pigs. *Animal* **2011**, *5*, 1283–1289.
7. Duarte, D.A.; Schroyen, M.; Mota, R.R.; Vanderick, S.; Gengler, N. Recent genetic advances on boar taint reduction as an alternative to castration: A Review. *J. Appl. Genet.* **2021**, *62*, 137–150.
8. Bonneau, M. EU research programme on boar taint: Summary of the results obtained so far and preliminary conclusions. In *Boar Taint in Entire Male Pigs*; Bonneau, M.; Lundström, M.; Malmfors, B., Eds.; Wageningen Academic Publishers, Wageningen, The Netherlands, 1997.
9. Bonneau, M.; Weiler, U. Pros and Cons of Alternatives to Piglet Castration: Welfare, Boar Taint, and Other Meat Quality Traits. *Animals* **2019**, *9*, 884.
10. Squires, E.J.; Bone, C.; Cameron, J. Pork Production with Entire Males: Directions for Control of Boar Taint. *Animals* **2020**, *10*, 1665.
11. Xue, J.; Dial, G.D.; Holton, E.E.; Vickers, Z.; Squires, E.J.; Lou, Y.; Godbout, D.; Morel, N. Breed differences in boar taint: Relationship between tissue levels boar taint compounds and sensory analysis of taint. *J. Anim. Sci.* **1996**, *74*, 2170.
12. Čandek-Potokar, M.; Žlender, B.; Lefaucheur, L.; Bonneau, M. Effects of age and/or weight at slaughter on longissimus dorsi muscle: Biochemical traits and sensory quality in pigs. *Meat Sci.* **1998**, *48*, 287–300.
13. Prosciutto di Parma (Parma Ham). Protected Designation of Origin (Specifications and Dossier Pursuant to Article 4 of Council Regulation EEC no. 2081/92 dated 14 July 1992). Available online: [https://www.prosciuttodiparma.com/wp-content/uploads/2019/07/Parma\\_Ham\\_Specifications\\_Disciplinare\\_Consolidato\\_Nov\\_13.pdf](https://www.prosciuttodiparma.com/wp-content/uploads/2019/07/Parma_Ham_Specifications_Disciplinare_Consolidato_Nov_13.pdf) (accessed on 12 March 2023).
14. Consorzio del Prosciutto di San Daniele. Production Specification for the Protected Designation of Origin “Prosciutto di San Daniele”. 2007. Available online: <https://www.prosciuttosandaniele.it/content/uploads/2018/05/Production-Specifications-SAN-DANIELE.pdf> (accessed on 6 January 2022).
15. Ducro-Steeverink, D. Selection against boar taint: A simulation study. *Acta Vet. Scand.* **2006**, *48*, 2006.
16. Mathur, P.K.; ten Napel, J.; Bloemhof, S.; Heres, L.; Knol, E.F.; Mulder, H.A. A human nose scoring system for boar taint and its relationship with androstenone and skatole. *Meat Sci.* **2012**, *91*, 414–422.
17. Windig, J.J.; Mulder, H.A.; ten Napel, J.; Knol, E.F.; Mathur, P.K.; Crump, R.E. Genetic parameters for androstenone, skatole, indole, and human nose scores as measures of boar taint and their relationship with finishing traits. *J. Anim. Sci.* **2012**, *90*, 2120–2129.
18. Rostellato, R.; Bonfatti, V.; Larzul, C.; Bidanel, J.P.; Carnier, P. Estimates of genetic parameters for content of boar taint compounds in adipose tissue of intact males at 160 and 220 days of age. *J. Anim. Sci.* **2015**, *93*, 4267–4276.
19. de Campos, C.F.; Lopes, M.S.; e Silva, F.F.; Veroneze, R.; Knol, E.F.; Sávio Lopes, P.; Guimarães, S.E.F. Genomic selection for boar taint compounds and carcass traits in a commercial pig population. *Livest. Sci.* **2015**, *174*, 10–17.

20. Larzul, C. How to improve meat quality and welfare in entire male pigs by genetics. *Animals* **2021**, *11*, 699.
21. Dugué, C.; Prunier, A.; Mercat, M.; Monziols, M.; Blanchet, B.; Larzul, C. Genetic determinism of boar taint and relationship with growth traits, meat quality and lesions. *Animal* **2020**, *14*, 1333–1341.
22. Grindflek, E.; Meuwissen, T.H.; Aasmundstad, T.; Hamland, H.; Hansen, M.H.; Nome, T.; Kent, M.; Torjesen, P.; Lien, S. Revealing genetic relationships between compounds affecting boar taint and reproduction in pigs. *J. Anim. Sci.* **2011**, *89*, 680–692.
23. Mathur, P.K.; ten Napel, J.; Crump, R.E.; Mulder, H.A.; Knol, E.F. Genetic relationship between boar taint compounds, human nose scores, and reproduction traits in pigs. *J. Anim. Sci.* **2013**, *91*, 4080–4089.
24. Parois, S.P.; Prunier, A.; Mercat, M.J.; Merlot, E.; Larzul, C. Genetic relationships between measures of sexual development, boar taint, health, and aggressiveness in pigs. *J. Anim. Sci.* **2015**, *93*, 3749–3758.
25. Brinke, I.; Große-Brinkhaus, C.; Roth, K.; Pröll-Cornelissen, M.J.; Henne, H.; Schellander, K.; Tholen, E. Genomic background and genetic relationships between boar taint and fertility traits in German Landrace and large white. *BMC Genet.* **2020**, *21*, 61.
26. Lervik, S.; Oskam, I.; Krogenæs, A.; Andresen, Ø.; Dahl, E.; Haga, H.A.; Tajet, H.; Olsaker, I.; Ropstad, E. Androstenone and testosterone levels and testicular morphology of Duroc boars related to estimated breeding value for androstenone. *Theriogenology* **2013**, *79*, 986–994.
27. Meuwissen, T.; Hayes, B.; Goddard, M. Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* **2016**, *6*, 6–14.
28. Botelho, M.E.; Lopes, M.S.; Mathur, P.K.; Knol, E.F.; e Silva, F.F.; Lopes, P.S.; Guimarães, S.E.; Marques, D.B.D.; Veroneze, R. Weighted genome-wide association study reveals new candidate genes related to boar taint compounds 1. *Livest. Sci.* **2022**, *257*, 104845.
29. Große-Brinkhaus, C.; Storck, L.C.; Frieden, L.; Neuhoﬀ, C.; Schellander, K.; Looft, C.; Tholen, E. Genome-wide association analyses for boar taint components and testicular traits revealed regions having pleiotropic effects. *BMC Genet.* **2015**, *16*, 36.
30. Meuwissen, T.; Hayes, B.; MacLeod, I.; Goddard, M. Identification of genomic variants causing variation in quantitative traits: A review. *Agriculture* **2022**, *12*, 1713.
31. Verardo, L.L.; Silva, F.F.; Varona, L.; Resende, M.D.; Bastiaansen, J.W.; Lopes, P.S.; Guimarães, S.E. Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. *J. Appl. Genet.* **2014**, *56*, 123–132.
32. Verardo, L.L.; Lopes, M.S.; Wijga, S.; Madsen, O.; Silva, F.F.; Groenen, M.A.; Knol, E.F.; Lopes, P.S.; Guimarães, S.E. After genome-wide association studies: Gene Networks elucidating candidate genes divergences for number of teats across two pig populations. *J. Anim. Sci.* **2016**, *94*, 1446–1458.
33. Engesser, D.J. Alternatives for Boar Taint Reduction and Elimination Besides Surgical Castration and Destroying Testicular Tissue. Ph.D. Thesis, University of Leipzig, Leipzig, Germany, 2015. Available online: <https://d-nb.info/1239566905/34> (accessed on 12 March 2023).
34. Moe, M.; Grindflek, E.; Doran, O. Expression of 3 $\beta$ -hydroxysteroid dehydrogenase, cytochrome P450-C17, and sulfotransferase 2B1 proteins in liver and testis of pigs of two breeds: Relationship with adipose tissue androstenone concentration. *J. Anim. Sci.* **2007**, *85*, 2924–2931.
35. Moe, M.; Lien, S.; Bendixen, C.; Hedegaard, J.; Hornshøj, H.; Berget, I.; Meuwissen, T.H.E.; Grindflek, E. Gene expression profiles in liver of pigs with extreme high and low levels of androstenone. *BMC Vet. Res.* **2008**, *4*, 29.
36. Grindflek, E.; Berget, I.; Moe, M.; Oeth, P.; Lien, S. Transcript profiling of candidate genes in testis of pigs exhibiting large differences in androstenone levels. *BMC Genet.* **2010**, *11*, 4.
37. Gunawan, A.; Sahadevan, S.; Neuhoﬀ, C.; Große-Brinkhaus, C.; Gad, A.; Frieden, L.; Tesfaye, D.; Tholen, E.; Looft, C.; Uddin, M.J.; et al. RNA deep sequencing reveals novel candidate genes and polymorphisms in boar testis and liver tissues with divergent androstenone levels. *PLoS ONE* **2013**, *8*, e63259.
38. Kelly, D.M.; Jones, T.H. Testosterone: A metabolic hormone in health and disease. *J. Endocrinol.* **2013**, *217*, R25–R45.
39. Baes, C.; Mattei, S.; Luther, H.; Ampuero, S.; Sidler, X.; Bee, G.; Spring, P.; Hofer, A. A performance test for boar taint compounds in live boars. *Animal* **2013**, *7*, 714–720.
40. Andresen, Ø. Concentrations of fat and plasma 5 $\alpha$ -androstenone and plasma testosterone in boars selected for rate of body weight gain and thickness of back fat during growth, sexual maturation and after mating. *J. Reprod. Fertil.* **1976**, *48*, 51–59.
41. Babol, J.; Squires, E.J. Liver metabolic activities of intact male pigs injected with skatole. *Can. J. Anim. Sci.* **1999**, *79*, 549–552.
42. Sinclair, P.A.; Squires, E.J.; Raeside, J.I. Early postnatal plasma concentrations of testicular steroid hormones, pubertal development, and carcass leanness as potential indicators of boar taint in market weight intact male pigs. *J. Anim. Sci.* **2001**, *279*, 1868–1876.
43. Boschi, E.; Faggion, S.; Mondin, C.; Carnier, P.; Bonfatti, V. Concentrations of boar taint compounds are weakly associated with sexual behavior of young boars. *Animals* **2022**, *12*, 1499.
44. Sargolzaei, M.; Chesnais, J.P.; Schenkel, F.S. A new approach for efficient genotype imputation using information from relatives. *BMC Genom.* **2014**, *15*, 478.
45. Misztal, I.; Tsuruta, S.; Lourenco, D.; Masuda, Y.; Aguilar, I.; Legarra, A.; Vitezica, Z. Manual for BLUPf90 Family of Programs. 2018. Available online: [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all7.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf) (accessed on 20 November 2022).
46. VanRaden, P.M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423.
47. Steiger, J.H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **1980**, *87*, 245–251.
48. Gualdrón Duarte, J.L.; Cantet, R.J.C.; Bates, R.O.; Ernst, C.W.; Raney, N.E.; Steibel, J.P. Rapid screening for phenotype-genotype associations by linear transformations of genomic evaluations. *BMC Bioinform.* **2014**, *15*, 246.



49. Bernal Rubio, Y.L.; Gualdrón Duarte, J.L.; Bates, R.O.; Ernst, C.W.; Nonneman, D.; Rohrer, G.A.; King, A.; Shackelford, S.D.; Wheeler, T.L.; Cantet, R.J.; et al. Meta-analysis of genome-wide association from Genomic Prediction Models. *Anim. Genet.* **2015**, *47*, 36–48.
50. Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Muir, W.M. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* **2012**, *94*, 73–83.
51. Aguilar, I.; Legarra, A.; Cardoso, F.; Masuda, Y.; Lourenco, D.; Misztal, I. Frequentist P-values for large-scale single step genome-wide association, with an application to birth weight in American Angus cattle. *Genet. Sel. Evol.* **2019**, *51*, 28.
52. Veroneze, R.; Bastiaansen, J.W.M.; Knol, E.F.; Guimarães, S.E.F.; Silva, F.F.; Harlizius, B.; Lopes, M.S.; Lopes, P.S. Linkage disequilibrium patterns and persistence of phase in purebred and crossbred pig (*Sus scrofa*) populations. *BMC Genet.* **2014**, *15*, 126.
53. Turner, S.D. qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. *J. Open Source Softw.* **2018**, *3*, 731.
54. Lemos, M.V.; Chiaia, H.L.; Berton, M.P.; Feitosa, F.L.; Aboujaoud, C.; Camargo, G.M.; Baldi, F. Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nellore cattle using the single step procedure. *BMC Genom.* **2016**, *17*, 213.
55. Marques, D.B.; Bastiaansen, J.W.; Broekhuijse, M.L.; Lopes, M.S.; Knol, E.F.; Harlizius, B.; Guimarães, S.E.; Silva, F.F.; Lopes, P.S. Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. *Genet. Sel. Evol.* **2018**, *50*, 40.
56. Sollero, B.P.; Junqueira, V.S.; Gomes, C.C.; Caetano, A.R.; Cardoso, F.F. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. *Genet. Sel. Evol.* **2017**, *49*, 49.
57. Desmoulin, B.; Bonneau, M.; Frouin, A.; Bidard, J.P. Consumer Testing of Pork and Processed Meat from Boars: The Influence of Fat Androstene Level. *Livest. Prod. Sci.* **1982**, *9*, 707–715.
58. Walstra, P.; Claudi-Magnussen, C.; Chevillon, P.; von Seth, G.; Diestre, A.; Matthews, K.R.; Homer, D.B.; Bonneau, M. An International Study on the Importance of Androstene and Skatole for Boar Taint: Levels of Androstene and Skatole by Country and Season. *Livest. Prod. Sci.* **1999**, *62*, 15–28.
59. Lukić, B.; Pong-Wong, R.; Rowe, S.J.; de Koning, D.J.; Velandar, I.; Haley, C.S.; Archibald, A.L.; Woolliams, J.A. Efficiency of genomic prediction for boar taint reduction in Danish Landrace pigs. *Anim. Genet.* **2015**, *46*, 607–616.
60. Botelho, M.E.; Lopes, M.S.; Mathur, P.K.; Knol, E.F.; Guimarães, S.E.; Marques, D.B.; Lopes, P.S.; Silva, F.F.; Veroneze, R. Applying an association weight matrix in weighted genomic prediction of boar taint compounds. *J. Anim. Breed. Genet.* **2020**, *138*, 442–453.
61. Gao, X.; Becker, L.C.; Becker, D.M.; Starmer, J.D.; Province, M.A. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* **2009**, *34*, 10–105.
62. Quintanilla, R.; Demeure, O.; Bidanel, J.P.; Milan, D.; Iannuccelli, N.; Amigues, Y.; Gruand, J.; Renard, C.; Chevalet, C.; Bonneau, M. Detection of quantitative trait loci for fat androstene levels in pigs. *J. Anim. Sci.* **2003**, *81*, 385–394.
63. Lee, G.J.; Archibald, A.L.; Law, A.S.; Lloyd, S.; Wood, J.; Haley, C.S. Detection of quantitative trait loci for androstene, Skatole and boar taint in a cross between large white and Meishan Pigs. *Anim. Genet.* **2005**, *36*, 14–22.
64. Varona, L.; Vidal, O.; Quintanilla, R.; Gil, M.; Sánchez A.; Folch, J.M.; Hortos, M.; Rius, M.A.; Amills, M.; Noguera, J.L. Bayesian analysis of quantitative trait loci for boar taint in a landrace outbred population. *J. Anim. Sci.* **2005**, *83*, 301–307.
65. Duijvesteijn, N.; Knol, E.F.; Merks, J.W.M.; Crooijmans, R.P.M.A.; Groenen, M.A.M.; Bovenhuis, H.; Harlizius, B. A genome-wide association study on Androstene levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genet.* **2010**, *11*, 42.
66. Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Fernando, R.L.; Vitezica, Z.; Okimoto, R.; Wing, T.; Hawken, R.; Muir, W.M. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. *Front. Genet.* **2014**, *5*, 134.
67. Grindflek, E.; Szyda, J.; Liu, Z.; Lien, S. Detection of quantitative trait loci for meat quality in a commercial slaughter pig cross. *Mamm. Genome* **2001**, *12*, 299–304.
68. Duijvesteijn, N.; Knol, E.F.; Bijma, P. Boar taint in entire male pigs: A genomewide association study for direct and indirect genetic effects on Androstene. *J. Anim. Sci.* **2014**, *92*, 4319–4328.
69. Milan, D.; Bidanel, J.-P.; Le Roy, P.; Chevalet, C.; Woloszyn, N.; Caritez, J.C.; Ollivier, L. Current status of QTL detection in Large White x Meishan crosses in France. In Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia, 11–16 January 1998; pp. 414–417.
70. Bidanel, J.-P.; Riquet, J.; Gruand, J.; Squires, J.; Bonneau, M.; Milan, D. Detection of quantitative trait loci for skatole and indole levels in Meishan x Large White F2 pigs. In Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13–18 August 2006; pp. 6–17.
71. Morozumi, T.; Sumantri, C.; Nakajima, E.; Kobayashi, E.; Asano, A.; Oishi, T.; Mitsuhashi, T.; Watanabe, T.; Hamasima, N. Three Types of Polymorphisms in Exon 14 in Porcine Mx1 Gene. *Biochem. Genet.* **2001**, *39*, 251–260.
72. Gregersen, V.R.; Conley, L.N.; Sørensen, K.K.; Guldbrandtsen, B.; Velandar, I.H.; Bendixen, C. Genome-wide association scan and phased haplotype construction for quantitative trait loci affecting boar taint in three pig breeds. *BMC Genom.* **2012**, *13*, 22.
73. Rowe, S.J.; Karacaören, B.; de Koning, D.-J.; Lukic, B.; Hastings-Clark, N.; Velandar, I.; Haley, C.S.; Archibald, A.L. Analysis of the genetics of Boar Taint reveals both single SNPs and regional effects. *BMC Genom.* **2014**, *15*, 424.

74. Zhang, H.; Na, W.; Zhang, H.-L.; Wang, N.; Du, Z.-Q.; Wang, S.-Z.; Wang, Z.-P.; Zhang, Z.; Li, H. TCF21 is related to testis growth and development in broiler chickens. *Genet. Sel. Evol.* **2017**, *49*, 25.
75. Bhandari, R.K.; Sadler-Rigglesman, I.; Clement, T.M.; Skinner, M.K. Basic helix-loop-helix transcription factor TCF21 is a downstream target of the male sex determining Gene Sry. *PLoS ONE* **2011**, *6*, e19935.
76. Munier, A.; Feral, C.; Milon, L.; Pinon, V.P.-B.; Gyapay, G.; Capeau, J.; Guellaen, G.; Lacombe, M.-L. A new human nm23homo-*logue* (nm23-h5) specifically expressed in testis germinal cells. *FEBS Lett.* **1998**, *434*, 289–294.
77. Choi, Y.-J.; Cho, S.-K.; Hwang, K.-C.; Park, C.; Kim, J.-H.; Park, S.-B.; Hwang, S.; Kim, J.-H. NM23-M5 mediates round and elongated spermatid survival by regulating GPX-5 levels. *FEBS Lett.* **2009**, *583*, 1292–1298.
78. Estienne, M.J.; Harper, A.F.; Knight, J.W.; Barb, C.R.; Rampacek, G.B. Sexual behavior after treatment with prostaglandin-F2A in boars with suppressed concentrations of gonadal steroids. *Appl. Anim. Behav. Sci.* **2004**, *89*, 53–57.
79. Hemsworth, P.H.; Tilbrook, A.J. Sexual behavior of male pigs. *Horm. Behav.* **2007**, *52*, 39–44.
80. Hammoud, S.S.; Purwar, J.; Pflueger, C.; Cairns, B.R.; Carrell, D.T. Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. *Fertil. Steril.* **2010**, *94*, 1728–1733.
81. Rajender, S.; Avery, K.; Agarwal, A. Epigenetics, spermatogenesis and male infertility. *Mutat. Res. Rev. Mutat. Res.* **2011**, *727*, 62–71.
82. Wang, W.; Tang, Y.; Ni, L.; Kim, E.; Jongwutiwes, T.; Hourvitz, A.; Zhang, R.; Xiong, H.; Liu, H.-C.; Rosenwaks, Z. Overexpression of uromodulin-like1 accelerates follicle depletion and subsequent ovarian degeneration. *Cell Death Dis.* **2012**, *3*, e433.
83. Russo, V.; Fontanesi, L.; Scotti, E.; Beretti, F.; Davoli, R.; Nanni Costa, L.; Virgili, R.; Buttazzoni, L. Single nucleotide polymorphisms in several porcine cathepsin genes are associated with growth, carcass, and production traits in Italian large white pigs. *J. Anim. Sci.* **2008**, *86*, 3300–3314.
84. Fontanesi, L.; Speroni, C.; Buttazzoni, L.; Scotti, E.; Dall'Olio, S.; Nanni Costa, L.; Davoli, R.; Russo, V. The insulin-like growth factor 2 (IGF2) gene intron3-g.3072g>A polymorphism is not the only sus scrofa chromosome 2p mutation affecting meat production and carcass traits in pigs: Evidence from the effects of a cathepsin D (CTSD) gene polymorphism. *J. Anim. Sci.* **2010**, *88*, 2235–2245.
85. Virgili, R.; Parolari, G.; Schivazappa, C.; Bordini, C.S.; Borri, M. Sensory and texture quality of dry-cured ham as affected by endogenous cathepsin B activity and muscle composition. *J. Food Sci.* **1995**, *60*, 1183–1186.
86. Kaltnekar, T.; Škrlep, M.; Batorek Lukač, N.; Tomažin, U.; Prevolnik Povše, M.; Labussiere, E.; Čandek-Potokar, M. Effects of Salting Duration and Boar Taint Level on Quality of Dry-Cured Hams. In Proceedings of the 24th International Symposium “Animal Science Days”, Ptuj, Slovenia, 21–23 September 2016; pp. 132–137.
87. Xu, Z.; Wei, G.; Chepelev, I.; Zhao, K.; Felsenfeld, G. Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. *Nat. Struct. Mol. Biol.* **2011**, *18*, 372–378.
88. Moghadam, P.K.; Jackson, M.B. The functional significance of synaptotagmin diversity in neuroendocrine secretion. *Front. Endocrinol.* **2013**, *4*, 124.
89. Xu, Z.Y.; Yang, H.; Xiong, Y.Z.; Deng, C.Y.; Li, F.E.; Lei, M.G.; Zuo, B. Identification of three novel SNPs and association with carcass traits in porcine TNNI1 and TNNI2. *Mol. Biol. Rep.* **2010**, *37*, 3609–3613.
90. Yang, H.; Xu, Z.Y.; Lei, M.G.; Li, F.E.; Deng, C.Y.; Xiong, Y.Z.; Zuo, B. Association of 3 polymorphisms in porcine troponin I genes (TNNI1 and TNNI2) with meat quality traits. *J. Appl. Genet.* **2010**, *51*, 51–57.
91. Oguri, Y.; Shinoda, K.; Kim, H.; Alba, D.L.; Bolus, W.R.; Wang, Q.; Kajimura, S. CD81 controls beige fat progenitor cell growth and energy balance via Fak Signaling. *Cell* **2020**, *182*, 563–577.
92. Boyle, K.B.; Hadaschik, D.; Virtue, S.; Cawthorn, W.P.; Ridley, S.H.; O’Rahilly, S.; Siddle, K. The transcription factors EGR1 and EGR2 have opposing influences on adipocyte differentiation. *Cell Death Differ.* **2009**, *16*, 782–789.
93. Lee, K.-T.; Byun, M.-J.; Kang, K.-S.; Park, E.-W.; Lee, S.-H.; Cho, S.; Kim, H. Neuronal genes for subcutaneous fat thickness in human and pig are identified by local genomic sequencing and combined SNP association study. *PLoS ONE* **2011**, *6*, e16356.
94. Xue, Y.; Li, C.; Duan, D.; Wang, M.; Han, X.; Wang, K.; Qiao, R.; Li, X.J.; Li, X.L. Genome-Wide association studies for growth-related traits in a crossbreed pig population. *Anim. Genet.* **2020**, *52*, 217–222.
95. Kim, E.-S.; Ros-Freixedes, R.; Pena, R.N.; Baas, T.J.; Estany, J.; Rothschild, M.F. Identification of signatures of selection for intramuscular fat and backfat thickness in two Duroc populations. *J. Anim. Sci.* **2015**, *93*, 3292–3302.
96. Luo, H.F.; Wei, H.K.; Huang, F.R.; Zhou, Z.; Jiang, S.W.; Peng, J. The effect of linseed on intramuscular fat content and adipogenesis related genes in skeletal muscle of pigs. *Lipids* **2009**, *44*, 999.
97. Little, H.C.; Rodriguez, S.; Lei, X.; Tan, S.Y.; Stewart, A.N.; Sahagun, A.; Sarver, D.C.; William Wong, G. Myonectin deletion promotes adipose fat storage and reduces liver steatosis. *FASEB J.* **2019**, *33*, 8666–8687.
98. Faggion, S.; Carnier, P.; Bonfatti, V. Genetic correlations between boar taint compound concentrations in fat of purebred boars and production and ham quality traits in crossbred heavy pigs. *Animals* **2023**, *13*, 2445. <https://doi.org/10.3390/ani13152445>.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.