Livestock Science 302 (2025) 105844

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Exploratory study of the potential use of a novel pooled phenotype to select against boar taint

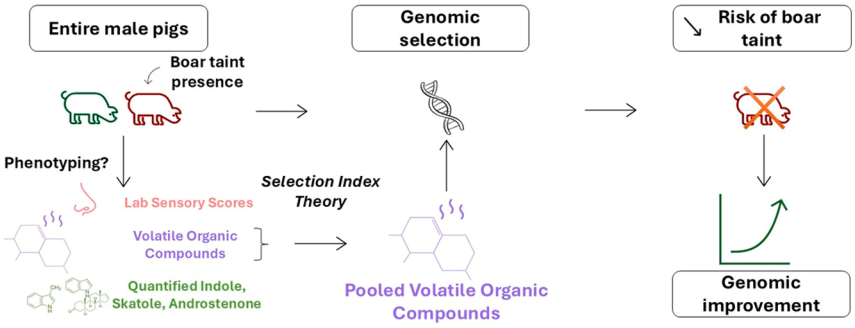
Alice Markey [a,*], Anaïs Rodrigues [b], Pierre-Hugues Stefanuto [b], Jean-François Focant [b], Anne-Catherine Huet [c], José Wavreille [d], Katrien Wijnrocx [a], Nicolas Gengler [a]

[a] TERRA Teaching and Research Center, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium
[b] University of Liège, Organic and Biological Analytical Chemistry Group, MolSys Research Unit, 4000 Liège, Belgium
[c] CER Groupe, Analytical Laboratory, 6900 Marloie, Belgium
[d] Animal Production Unit, Walloon Agricultural Research Centre, 5030 Gembloux, Belgium

HIGHLIGHTS

- The genetic potential of recently found boar taint compounds was investigated.
- Genetic contributions were pooled into genetic index finding concise phenotype.
- Indexes resulted in high correlations to human perception.
- Research supported interest in continued efforts in high-throughput phenotyping.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Androstenone
Genetic prediction
Indole
Lab sensory score
Skatole
Volatile organic compounds

ABSTRACT

Meat quality traits are economically important in pig production. Breeding strategies can help prevent meat defects such as boar taint, usually characterized by quantified indole, skatole and androstenone (ISA) in back fat. This exploratory study investigated the genetic potential of a novel boar taint phenotype, pooling volatile organic compounds (VOCs), which were recently identified as phenotypically discriminant. Fat samples were collected from 1272 Pietrain × Landrace crossbred boars. Phenotypes for boar taint on these samples were: lab sensory score (LSS; $n = 1269$), ISA quantification ($n = 308$), and VOC profiles ($n = 127$). Given the limited amount of data, a selection index-based approach was used to pool traits in trait groups, ISA and VOC, considering LSS as reference trait. (Co)variance components were estimated with a full multi-trait model, and index equations were adjusted to account for uncertainty in estimated parameters. Index coefficients were then applied to ISA and VOC phenotypes to generate two pooled phenotypes, ISA and VOC indices. Estimates from the 3-trait model (LSS, ISA index and VOC index) confirmed high expected correlations with LSS. Genetic parameter estimates showed higher significance demonstrating the interest of pooling multiple partially informative traits together. Moreover, using the VOC index would generate a higher expected correlated genetic response in LSS (192 %) than the ISA index (160 %) compared to the direct response when using only LSS. Despite limited data, this exploratory

* Corresponding author.
E-mail address: alice.markey@uliege.be (A. Markey).

https://doi.org/10.1016/j.livsci.2025.105844
Received 24 July 2025; Received in revised form 23 October 2025; Accepted 24 October 2025
Available online 30 October 2025

study showed the potential of this novel broad phenotype based on pooled VOCs to improve genetic selection for reduced boar taint risk, although further validation in larger populations is required.

1. Introduction

Uncastrated males, bred for welfare, growth and feed efficiency, may develop boar taint – an undesirable (i.e., faecal and urinary) flavour released during meat cooking. Boar taint can be assessed by sensory evaluation, performed at slaughterhouse or in the lab, by heating fat and detecting released volatile organic compounds (VOCs) (Burgeon et al., 2021a; Mörlein et al., 2014). This method, close but not identical to consumer's perception, enables to perceive the VOCs responsible for boar taint and their synergistic effects (Burgeon et al., 2021a). The perception of the boar taint phenotype has been differently characterized according to some studies. Some reported equivalent contributions of skatole (SKA) and androstenone (AND) (Berg et al., 1993; Hansson et al., 1980), while other stressed higher contributions from SKA relative to AND (Bejerholm and Barton-Garde, 1993; Bonneau et al., 2000). The contribution of both compounds to total perception was reported to range between 50 % to 66 % (Hansson et al., 1980; Bejerholm and Barton-Garde, 1993) and their existing interaction and influence on perception has been demonstrated by Mörlein et al. (2016a). These differences between studies can be explained because perception can be affected by concentrations, animal populations, sensory analysis methods, assessors and consumption context (Bonneau, 1993; Markey et al., 2025; Mörlein et al., 2014, 2016a, 2016b; Rius et al., 2005). Other recent studies (e.g., Burgeon et al., 2021a; Fischer et al., 2014; Gerlach et al., 2018; Markey et al., 2025; Mörlein et al., 2024), reported, among other compounds, that indole (IND) and 2-aminoacetophenone also contribute to faecal and foxy odours, respectively thereby adding complexity to the boar taint phenotype.

Genetic selection of pigs to reduce boar taint risk (i.e., its perception by humans) generally uses traits like quantified compounds (i.e., SKA, AND, and sometimes IND) (Haberland et al., 2014; Squires et al., 2020). This approach relies on the strong genetic correlations, often estimated in ranges between 0.65 and 0.90 (Haberland et al., 2014) between indole, skatole and androstenone (ISA), as these compounds are the main contributors to the boar taint, and the direct human perception, evaluated as lab sensory score (LSS). Additionally, the relatively high heritability of SKA (0.23–0.55) and AND (0.55–0.88), whilst LSS exhibit lower heritability (0.07–0.30), guarantee a highest response to selection (Duarte et al., 2021; Markey et al., 2025).

However, the genetic evaluation of boar taint should account for human sensory perception which may not be fully associated to ISA traits only. Therefore, relying on a correlated response from these traits only is suboptimal and may miss to cover the full human perception. If additional compounds contributing to boar taint should exist, the next steps are their identification and quantification. This process currently results in a limited number of observations, restricting the scope for comprehensive genetic studies. Despite these limitations, even small-scale or preliminary studies are valuable, because they can enhance the understanding of how these compounds relate to human perception of boar taint and clarify if more efforts for phenotyping are necessary. Indeed, quantification requires analytical methods such as chromatography (Hansen-Møller et al., 1994; Windig et al., 2012; Burgeon et al., 2021a, 2021b; Botelho et al., 2022) although accurate are often costly and time-consuming (Burgeon et al., 2021a), which further limit large-scale implementation.

In order to detect multiple informative traits for boar taint perception, Rodrigues et al. (2024) recently identified 13 VOCs in pig backfat discriminating boar taint using a classification model. They used extraction of heated fat head-space method coupled to gas chromatography and mass spectrometry (GC–MS) which enabled them to obtain profiles of VOCs. This type of analysis was already used for boar taint characterisation by, among others, Fischer et al. (2011), Verplanken et al. (2016), Burgeon et al. (2021b, 2023a, 2023b). This analysis can be fast as the device can be portable (Verplanken et al., 2016) and can be used to quantify boar taint compounds (e.g., Fischer et al., 2011) or to study VOC profiles (e.g., Burgeon et al., 2021b) with low limits of quantification and detection. However, to our knowledge, its potential application for genetic purposes for selection against boar taint has not yet been exploited.

This exploratory study evaluated the potential of boar taint-related compounds identified by Rodrigues et al. (2024) in the genetic improvement context. These VOCs included volatilized forms of IND, SKA, AND and ten additional compounds discriminant in sensory analysis. The objective of this exploratory study was to assess the potential use of VOCs in the context of genetic selection against boar taint and their relationships with other traits: LSS, representing the boar taint phenotype reference trait, and ISA quantification. We studied their potential using two approaches: directly using VOCs individually and as, alternatively, as pooled VOCs according to their contribution to LSS. In the first approach, we considered individually each trait (LSS, ISA quantification and the 13 VOCs) and we estimated (co)variance components in a 17-trait model. In the second approach, we used results from the first part of this study to consider pooled ISA quantification and pooled VOCs to predict the boar taint phenotype. Respective coefficients were calculated based on selection index (SI) theory to pool weighted compounds according to their relative contribution in genetically explaining LSS, the reference trait. Then, (co)variances components were estimated from a 3-trait model including LSS and the two boar taint indices. Finally, the genetic relevance of pooled VOC-based phenotype was evaluated for its effectiveness in selection against boar taint, by comparing expected genetic responses to direct LSS and pooled ISA-based phenotype respectively.

2. Material and methods

This study was carried out as part of a larger research project on the development of selection tools against boar taint of Walloon Pietrain boars to explore new opportunities for phenotyping. The Pietrain × Landrace crossbred animals were born on the same farm, reared in 14 batches of 100 animals and each batch was then approximately split between two farms for fattening. Animals in each batch were slaughtered together at around 180 days. Each cohort of approximately 50 animals per batch and per farm represented a given contemporary group. For practical reasons, two contemporary groups were not reared resulting in 26 contemporary groups. No selection was applied on animals and conditions were similar to commercial farms. The pedigree of these animals and ancestors was registered, auricular biopsy was performed for genotype analysis, and back fats were collected at culling, allowing to obtain several phenotypes of boar taint. The phenotypic and the genotypic data in this study were obtained as part of the station progeny testing program of Walloon Pietrain boars and the project NoWallOdor to develop a routine boar taint genetic evaluation. The Ethics Committee of the University of Liege approved the protocol for animal raising and handling (Protocols codes #2307 #1861; approved on 21 December 2020).

2.1. Phenotypic data

Data was based on neck back fat samples from 1272 non-castrated boars slaughtered between 2020 and 2023. Back fat samples were collected on the slaughter line and stored at −20 °C until analyses. The back fats were divided into three samples, weighing 1 g, 50 g and 5 g

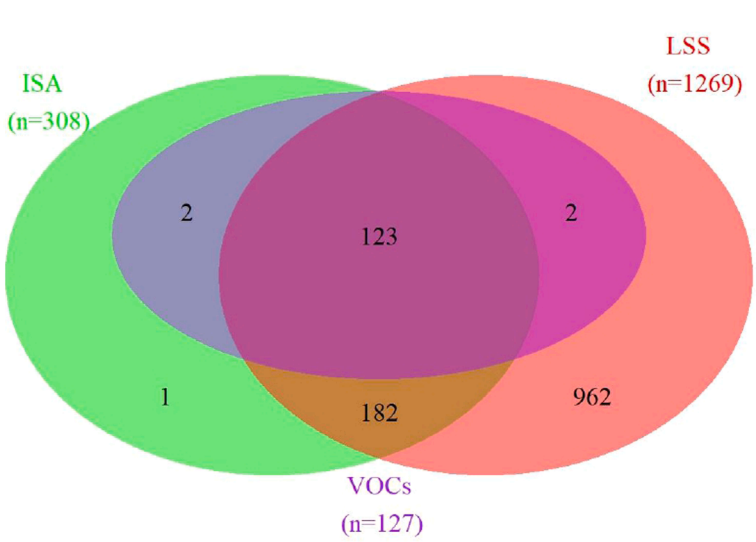
respectively for the following analyses. These samples were analysed for boar taint by three types of analyses resulting in three boar taint phenotypes: The LSS, the ISA quantification, and the VOCs. During data curation, following rules were applied: no duplicates were kept, only Piétrain × Landrace crossbred boars with at least one phenotype, known pedigree, not slaughtered under/overage (24–32 weeks) were used, each batch slaughtered at a common age with average weight of 119±16 kg. Animals with at least one phenotype were kept. Final dataset of phenotypes ($n = 1272$) consisted of LSS ($n = 1269$), ISA quantification ($n = 308$) and 13 VOCs ($n = 127$) data. In order to visualize the complex data structure, data overlap is presented in Fig. 1(a) as a Venn diagram. The following paragraphs give more details about used analytical methods.

2.1.1. Sensory analysis

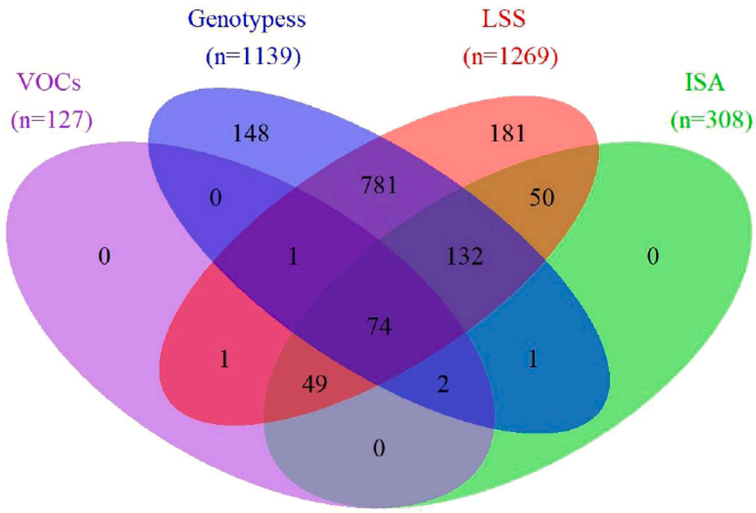
Sensory analyses were performed on almost all sampled animals, with a soldering iron by one expert, trained to detect boar taint levels, at laboratory, i.e., under controlled conditions (under a fume hood, organised with a 60 s interval between two successive analyses). This expert used a 3-scale classification of fat samples, resulting in LSS, as following: 0 (no odour), 1 (slight odour) and 2 (strong odour). Before each session of analyses, the trained assessor had to demonstrate the ability to detect these levels. Each batch of samples was assessed within the same session, as soon as possible after slaughter.

2.1.2. Quantification of indole, skatole and androstenone

Samples were selected throughout the batches and farms with a minimum of 7 quantifications by contemporary group. Resources were limited, therefore special attention was paid to ensure obtaining measurements of offspring for all evaluated sires applying an algorithm presented by Markey et al. (2023). Quantification of ISA (in ppb) was performed by ultra-high-performance liquid-chromatography tandem mass spectrometry (UHPLC-MS/MS) after an extraction from melted back fat with methanol by an external lab. Quantification was performed



LSS	ISA	VOCs	n
Yes	Yes	Yes	123
Yes	Yes	No	182
Yes	No	Yes	2
Yes	No	No	962
No	Yes	Yes	2
No	Yes	No	1



LSS	ISA	VOCs	Genotypes	n
Yes	Yes	Yes	Yes	74
Yes	Yes	Yes	No	49
Yes	Yes	No	Yes	132
Yes	Yes	No	No	50
Yes	No	Yes	Yes	1
Yes	No	Yes	No	1
Yes	No	No	Yes	781
Yes	No	No	No	181
No	Yes	Yes	Yes	2
No	Yes	No	Yes	1
No	No	No	Yes	148

Fig. 1. Availability data map after pre-processing of (a) phenotyped animals (b) genotyped across phenotyped animals.
Venn diagrams with associated tables of data combinations indicating the number of data available (Lab sensory score, indole, skatole and androstenone measured by ultra-high-performance liquid-chromatography tandem mass spectrometry, Volatile organic compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry, genotypes) after pre-processing across (a) phenotyped animals, (b) genotyped across phenotyped animals.
Abbreviations: LSS: Lab sensory score; ISA: Indole, skatole and androstenone measured by ultra-high-performance liquid-chromatography tandem mass spectrometry; VOCs: Volatile organic compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry.

with internal standards in all injections and was conducted by series, accompanied by a blank, a calibration curve and a control sample. More details about quantification and instrumentations are given by Rodrigues et al. (2024).

2.1.3. Volatile organic compounds analysis

Data for VOCs were based on measurements already used in the study of Rodrigues et al. (2024), with additional data curation applied (i.e., valid animal identification and eliminating repeated measurements). Back fat samples were randomly selected in six contemporary groups (i.e., three batches in two farms). Extraction was performed by heating back fat in a headspace and using a solid-phase microextraction fibre and the analysis was performed by comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry (HS-SPME-GCxGC-TOFMS). Instrumentation and other details were described in Rodrigues et al. (2024). The chromatograms were processed to obtain an exploitable dataset (i.e., pics area) for identifying compounds, carrying out statistical analyses, and building the classification model by cross-validation. Data for 13 VOCs ($n = 127$) identified by Rodrigues et al. (2024) as discriminant for boar taint were hereafter named as compounds C1 to C13: androstenone, skatole, 2-aminoacetophenone, indole, benzonitrile, 2,5-dimethyl-pyrazine, 1-methoxy-4-methyl-benzene, 2-propenyl-benzene, 2,3-dimethylphenyl isocyanate, benzenemethanol, biphenyl, pyrazine and 2-dodecanone.

2.2. Pedigree and genomic collection

The pedigree was extracted up to three generations ($n = 1842$) from the routine genetic Walloon evaluation system. The combined pedigree-genomic relationships (e.g., Christensen and Lund, 2010) were set up using genomic data ($n = 1139$) including 991 genotyped crossbred animals with at least one phenotype. Distribution of genotypes across available phenotypic data is shown in Fig. 1(b). All genotypes were available through the routine genetic Walloon evaluation system. Genotyping chips used were the GeneSeek-Neogen PorcineSNP80 BeadChip and the Illumina PorcineSNP60 BeadChip (Illumina, Inc., San Diego, CA). After quality control (Call rate > 0.90 and Minor Allele Frequency < 0.05) 40.928 single nucleotide polymorphisms (SNPs) remained.

2.3. Merging datasets

The different datasets (i.e., phenotypic datasets, pedigree and genotypes files) were merged based on common animal identification. Data availability and overlaps composition are shown in Fig. 1(a). SAS 9.4 software (SAS Institute, 2025) was used to prepare the file, transform and standardised the data and compile the basic statistics.

2.4. Genetic analyses

To investigate the genetic potential of VOCs, the general approach of this study was based on SI theory (Lin, 1978a, 1978b; Van Vleck, 1993), which is a multiple regression type approach where known information is used to predict unknown. The SI provides a synthetic value, called index, combining several traits weighted according to their relative importance. This allows individuals to be classified according to their overall value. The information to be predicted was LSS, the reference trait, based on known information from ISA traits and VOC traits. Therefore, ISA- and VOC-based predictions of LSS will be Best Linear Predictors of LSS creating alternative boar taint phenotypes (Van Vleck, 1993). In our context, we derived SI weights using genetic covariances among boar taint traits: LSS, ISA quantification and VOCs. Extending beyond classical use of SI, the derived coefficients were used to transform phenotypic data for two reasons. First, this allowed pooling of ISA quantification and VOCs, into new traits representing the boar taint phenotype, hereafter called ISA index and VOC index respectively. Moreover, this allowed the estimation of the final covariance components for LSS, ISA index and VOC index. All these covariance components were then used to assess the partitioning of genetic variance for LSS among traits and indices. Finally, literature is commonly used to obtain variance components for given traits when phenotypic data is limited or not available (e.g., as done by Haberland et al. (2014)). However, for VOC, to our knowledge, literature values are not yet available. Therefore, our study had to establish these (co)variance components itself despite the limited amount of data available. Moreover, by creating an index at a phenotypical level, the significance of estimates for unpooled and pooled traits could be compared.

2.4.1. Data preparation and statistical analysis

Boar taint is relatively infrequent in the pig population (5–10 %) even if it is a major problem of consumption and breeding (Rius et al., 2005; Sørensen et al., 2015). This results in an unbalanced population which may lead to statistical treatments less straightforward (Makridis et al., 2022). Moreover, data in this study are categorical (LSS) and continuous (ISA and VOC phenotypes). This is why, as previously applied by Markey et al. (2025), and similarly to Mathur et al. (2012), and based on Snell (1964) and Mujibi and Crews (2009), the use of lab sensory Snell-scores ($n = 1269$) was chosen as an approach to mitigate this problem. This method considers the frequency of each categorical data class, useful notably when a class is overrepresented as for boar taint. The LSS (0–2) were replaced by values representing the distribution of the three classes of observations. In practice, the cumulative frequencies at the boundaries of class 1 (i.e., the lower boundary represented by the frequency of class 0, and the upper boundary represented by the cumulative frequency of classes 0 and 1) were transformed into z-scores and then averaged. For the extreme classes, the 0 or 2 scores were replaced by approximated values that adjusted the first or last class in existing boundaries. This adjustment involved subtracting of adding the coefficient -ln(Pe)/Qe where Pe is the probability of a value below (for class 0) or above (for class 2) the boundary class and Qe is the relative proportion of each extreme class. For ISA measurements ($n = 308$), the $\log_{10}$-transformation was applied to improve normality (Tajet et al., 2006; Windig et al., 2012; Markey et al., 2025). Finally, all traits in the VOC dataset ($n = 127$), whose distributions resulting from statistical processing previously carried out by Rodrigues et al. (2024), were $\log_{10}$-transformed to approximate normality. All traits in the final dataset (17 traits; $n = 1272$) were standardised within trait to a mean of 0 and an observed standard deviation of 1.

2.4.2. Establishing initial genetic parameters using the full multi-trait model

Estimation of covariance components is inherently difficult with limited data. Therefore, many precautions were taken to achieve reliable results. As explained before, dataset was normalised and standardised to ensure uniformity and to achieve more stable convergence of the (co) variance component algorithms. Even if computation with several partial trait groups were tested leading to starting values, a full 17-trait model could be successfully run. The following generic multi-trait mixed model to estimate variance components, here after called full multi-trait model was used:

$$\mathbf{y} = \mathbf{Xf} + \mathbf{Zu} + \mathbf{e}$$

where $\mathbf{y}$ is the vector of traits × observations (all traits including LSS, ISA and VOCs for 1272 phenotyped animals); $\mathbf{X}$ and $\mathbf{Z}$ are the incidence matrices linking, respectively, fixed and random effects to records; $\mathbf{f}$ is the vector of fixed contemporary group effects with 26 classes of 14 batches in 2 farms and regression on metabolic slaughter weight for each trait; $\mathbf{u}$ is the vector of additive genomic random effects, $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G_0} \otimes$

H); **e** is the vector of residuals, **e** ~ N (**0**,$\mathbf{R_0} \otimes \mathbf{I}$); $\mathbf{G_0}$ and $\mathbf{R_0}$ are the elementary additive genetic and residual (co)variances matrices among traits, respectively; **I** is an identity matrix; **H** is the relationship matrix based on pedigree and genomic information, which inverse ($\mathbf{H^{-1}}$) was given by Aguilar et al. (2010, 2011) and reported by Botelho et al. (2022):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G_c}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

where, $\mathbf{A}^{-1}$ is the inverse of pedigree-based relationship matrix (**A**); $\mathbf{G_c}^{-1}$ is the inverse of the combined genomic relationship matrix ($\mathbf{G_c}$) with $\mathbf{G_c} = \alpha\mathbf{G} + \beta\mathbf{A}_{22}$ using by default $\alpha = 0.95$ and $\beta = 0.05$; $\mathbf{A}_{22}^{-1}$ is the inverse of pedigree-based relationship matrix for genotyped animals. The **G** matrix was calculated as described by VanRaden et al. (2009), Wang et al. (2012) and reported by Botelho et al. (2022):

$$\mathbf{G} = \mathbf{ZDZ}' \Big/ \left[2\sum_{i=1}^{m} p_i(1-p_i)\right]$$

wherein **Z** is a zero-centred matrix obtained by **Z** = **M** - **P**, **M** is a m × n (number of markers x number of animals) matrix, which specifies each individual genotype and **P** is a matrix with the allele frequencies expressed as a difference of 0.5 and multiplied by 2, i.e., the i column of **P** is given by 2 (p_i - 0.5); **D** is a diagonal matrix of SNP weights, here an identity matric; and p_i is the allele frequencies of the considered allele of the i^{th} marker. Variance components were estimated using expectation-maximisation and average information restricted maximum likelihood (REML) algorithms as implemented in the BLUPF90 family of programs (Misztal et al., 2014). Standard errors (SE) were based on Meyer and Houle (2013) and used to define significance of genetic parameters with Student's *t*-test, where p-values allowed to determine the significance threshold.

2.4.3. Regulation of initial genetic parameters

Although feasible, the parameters estimation method may produce unreliable results due to multicollinearity, as all variables were initially selected for their predictive potential on LSS. The presence of large SE led to modifications of the estimated covariance matrix, following Meyer and Kirkpatrick (2010). Rather than directly bending the matrix during estimation as suggested by these authors, a data-driven regularization by inflating the diagonals of the elementary genetic (co)variance matrix $\mathbf{G_0}$ among traits with their estimated SE was applied generating a modify matrix $\mathbf{G_0^*}$. This approach, conceptually Empirical Bayesian Ridge Regression, prevents overfitting and enhances stability by applying and L2 penalty. Unlike standard ridge regression, this method adapts to the data ensuring biologically meaningful variance adjustments while stabilizing estimated parameters and therefore estimation of SI coefficients. It differs from penalised regression as uncertainty-based regularization of covariance components was performed instead of direct shrinking regression coefficient. This led to reduced correlations among traits reducing multicollinearity issues. All matrix manipulations described in this section were performed using the program GNU Octave (Eaton et al., 2020).

2.4.4. Establishing selection indices and using them to create pooled predicted phenotypes

This section outlines the development and subsequent use of the SI to predict LSS. Indices were obtained and used in following steps: (1) Modifying ($\mathbf{G_0^*}$), the elementary genetic covariance matrix of traits (i.e., LSS, ISA and VOC traits), as explained in section 2.4.3.; (2) Obtaining the regression coefficients (i.e., required SI coefficients) of LSS on ISA traits and on VOC traits (Lin, 1978a): Regression coefficients ($\mathbf{b}_i$) were obtained from the previous elements by computing $\mathbf{b}_i' = \mathbf{c}_i'\left(\mathbf{G}_0^*\right)_i^{-1}$ where the vector $\mathbf{c}_i$ contains the genetic covariances between LSS and ISA and LSS and VOC traits while the matrix $\left(\mathbf{G}_0^*\right)_i$ contains the genetic covariances among ISA respectively VOC traits where i represents the computed index (Lin, 1978b); (3) Creating the pooled traits: From the $\mathbf{b}_i$ coefficients, a transformation matrix (**T**) was obtained and, for each animal, applied to the 17 traits to generate the pooled boar taint traits (i.e., ISA index and VOC index) representing predictions of LSS:

$$\begin{bmatrix} y_{LSS} \\ y_{ISA\ index} \\ y_{VOC\ index} \end{bmatrix} = \begin{bmatrix} 1_{LSS} & \mathbf{0}_{ISA} & \mathbf{0}_{VOC} \\ 0_{LSS} & \mathbf{b}_{ISA}' & \mathbf{0}_{VOC} \\ 0_{LSS} & \mathbf{0}_{ISA} & \mathbf{b}_{VOC}' \end{bmatrix} \begin{bmatrix} y_{LSS} \\ \mathbf{y_{ISA}} \\ \mathbf{y_{VOC}} \end{bmatrix} = \mathbf{T}\mathbf{y}_{LSS,\ ISA,\ VOC}$$

$$\text{where} \begin{bmatrix} y_{LSS} \\ y_{ISA\ index} \\ y_{VOC\ index} \end{bmatrix} = \mathbf{y}_{LSS,\ predicted\ LSS}$$

As traits were available by group of traits, the equation given above was modified for animals having only LSS ($n = 1269$), ISA ($n = 308$) or VOCs ($n = 127$) or any combination of those trait groups. The coefficients of **T** were optimised to genetically predict LSS from the ISA and VOC trait groups. By applying these coefficients to the phenotypic values, pooled predicted boar taint phenotypes were created, hereafter referred to as the ISA index and the VOC index.

2.4.5. Computing relative genetic contributions to LSS

Relative genetic contributions of each trait to LSS within their respective index were estimated using the following procedure: (1) Regression coefficients as those used in **T** were expressed for standardised genetic values; (2) The absolute values of these standardised coefficients were then normalised relative to their total size, resulting in relative contribution of each trait to the model; (3) By multiplying those coefficients by the total explained relative variance of LSS (i.e., coefficient of determination). This results in the relative genetic contributions of each trait to LSS.

2.4.6. Estimation of genetic parameters and genetic response for indexes and LSS

Covariance components estimation of the indices associated to LSS were performed with the same methods and the same model as before but with a reduced 3-trait model using the same procedure as section 2.4.2. The vector of traits × observations was composed of maximum three traits (LSS, ISA index and VOC index). The efficiency of indirect selection by selected trait (i.e., ISA or VOC index) compared to direct selection (by LSS) has been estimated according to the following standard formular based on SI-theory (Van Vleck, 1993):

$$\frac{CR_{LSS}}{R_{LSS}} = r_g\sqrt{\frac{h^2_{Selected\ trait}}{h^2_{LSS}}}$$

where CR_{LSS} is the correlated response of LSS by selecting for the selected trait (i.e., ISA or VOC indices), R_{LSS} is the direct response, r_g is the genetic correlation between traits, and h^2 are the heritabilities.

3. Results and discussion

3.1. Phenotypic data description

Basic characteristics of data are given in Table 1 and their overlap is shown in Fig. 1. The traits are shown before their standardisation based on the phenotypic mean, which was performed to optimise the use of multi-trait REML algorithms and their numerical stability.

3.2. Estimating initial genetic parameters among all traits

Needed covariances could be estimated jointly using the 17-trait model and sub-sampling to smaller matrices was not necessary.

Table 1
Descriptive statistics of the 17 transformed traits before standardisation.

Trait	N	Mean	Standard Deviation	Minimum	Maximum
Snell(LSS)	1269	0.27	0.66	0.02	2.71
Log_{10}(IND) (ppb)	308	1.39	0.71	0.39	3.03
Log_{10}(SKA) (ppb)	308	1.80	0.94	−0.06	3.47
Log_{10}(AND) (ppb)	308	3.00	0.60	0.70	3.91
Log_{10}(C1)	127	5.01	0.32	4.14	5.74
Log_{10}(C2)	127	4.55	0.34	3.72	5.27
Log_{10}(C3)	127	4.65	0.26	3.94	5.24
Log_{10}(C4)	127	4.94	0.34	3.62	5.71
Log_{10}(C5)	127	5.35	0.24	4.79	6.05
Log_{10}(C6)	127	5.05	0.35	4.36	6.24
Log_{10}(C7)	127	4.83	0.30	4.25	5.99
Log_{10}(C8)	127	4.61	0.26	3.76	5.66
Log_{10}(C9)	127	4.57	0.29	3.80	5.13
Log_{10}(C10)	127	4.89	0.22	4.38	5.40
Log_{10}(C11)	127	4.14	0.20	3.66	4.75
Log_{10}(C12)	127	5.91	0.37	3.86	6.50
Log_{10}(C13)	127	5.54	1.03	3.67	7.11

Abbreviations: LSS: Lab sensory score; IND, SKA and AND are compounds measured by ultra-high-performance liquid-chromatography tandem mass spectrometry; IND: Indole; SKA: Skatole; AND: Androstenone; C1-C13 are compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry; C1: Androstenone; C2: Skatole; C3: 2-Aminoacetophenone; C4: Indole; C5: Benzonitrile; C6: 2,5-Dimethyl-pyrazine; C7: 1-Methoxy-4-methyl-benzene; C8: 2-Propenyl-benzene; C9: 2,3-dimethylphenyl isocyanate; C10: Benzenemethanol; C11: Biphenyl; C12: Pyrazine; C13: 2-Dodecanone.

Observed convergence, presented in Fig. 2, showed a slow, but regular (Misztal, 2008), behaviour needing over 3000 rounds but eventually achieving a convergence of under 1^{-12}. There are several potential reasons that could explain this successful, slow but stable convergence. First, data was standardized creating the numerically most stable conditions. Furthermore, data had certainly large ranges of missing traits for a given record, but these were organized in three trait groups that showed strong genetic correlations inside the groups and across them (Table 2). Finally, the population structure of the dataset and how relationships were considered, was certainly helpful in achieving convergence. Indeed, this data came from a boar-testing scheme with batches carefully linked by connection sires across batches, but also with 78 % of animals with records being genotyped as shown in Fig. 1(b). Moreover, a large majority of sires and dams were also genotyped creating even more density in the combined relationship matrices.

Heritabilities, genetic and phenotypic correlations of the 17 traits and associated p-values (Table 2) were directly computed from estimated variance components $\mathbf{G_0}$ and $\mathbf{R_0}$. Heritabilities ranged from 0.16 to 0.66, genetic correlations from −0.90 to 0.97 and phenotypic correlations from −0.64 to 0.73. Out of the 17 heritabilities, most (11) were significant ($p < 0.05$) including very highly significant ($p < 0.001$) for two. Out of the 136 phenotypic correlations, 77 were significant ($p < 0.05$), most of which (40) were very highly significant ($p < 0.001$). Out of the 136 genetic correlations tested, the majority (129) were not significant, only three correlations had $p < 0.05$, and four showed a trend towards significance, i.e., $p < 0.1$. P-values associated to phenotypic correlations between traits were generally indicating significance, most likely due to common environmental effects. In contrast, since p-values for most of genetic correlations were high, these correlations were generally less significant, indicating a more limited and less certain shared genetic basis among the studied traits. However, strong evidence suggests that VOCs jointly contribute to the boar taint perception (Rodrigues et al., 2024) which motivated our development of a pooling method to study the genetic relationships of VOCs with other boar taint phenotypes considering them as a whole.

The heritability of LSS was estimated at 0.16 (low) while heritabilities for ISA traits were moderate to high (0.44–0.53) and at least highly significant ($p < 0.01$). The genetic correlations between ISA traits and LSS were medium to strong (0.59–0.82) and significant for IND and SKA ($p < 0.05$) but not for AND ($p < 0.1$) which showed only a trend. Phenotypic correlations between LSS and ISA quantification were moderate (0.25–0.36) and very highly significant ($p < 0.001$). These expected results as ISA traits are considered important as important components of boar taint perception. Results are in line with Windig et al. (2012) who found similar genetic parameters of boar taint phenotypes even though different methods were used. It appears that when selecting for a lower boar taint incidence, selection should mainly affect SKA and to a lesser extent the other compounds. Although LSS are less accurate, especially due to potential false negatives, they are widely available, obtained in a time-efficient manner (Burgeon et al., 2021a) and they may reflect more accurately final consumers' perceptions of boar taint. Finally, LSS can be useful for genetic selection against major compounds responsible for boar taint since this trait was found to be strongly correlated with ISA traits.

Heritability estimates of the 13 VOC traits (C1 to C13) were moderate to high (0.33–0.66), some of them were significant (7; $p < 0.05$), the others were showing a trend towards significance (1; $p < 0.1$) or

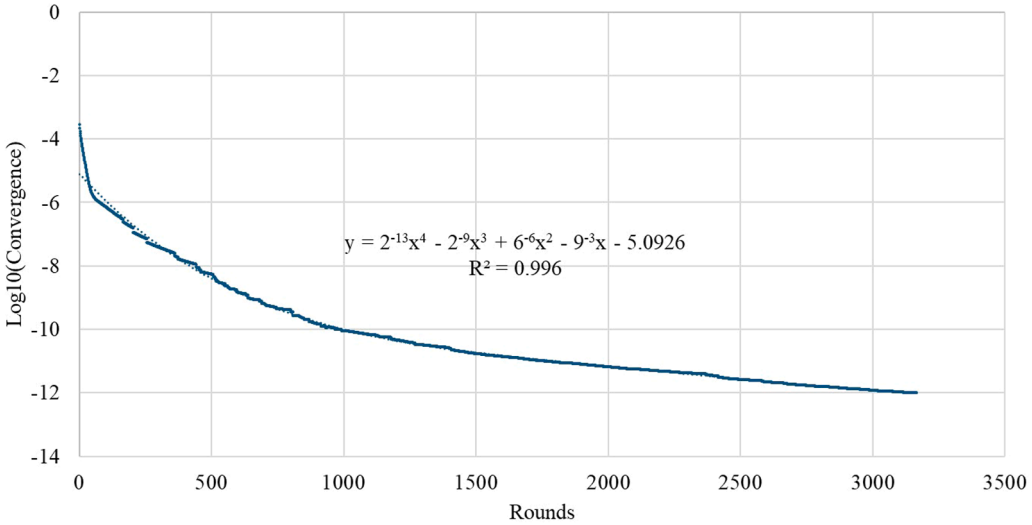


Fig. 2. Evolution of the convergence during variance component estimations by BLUPF90+.

Table 2
Heritabilities (on the diagonal), genetic (above) and phenotypic (below) correlations of lab sensory score, indole, skatole and androstenone measured by ultra-high-performance liquid-chromatography tandem mass spectrometry, and compounds (C1 to C13) measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry.

	LSS	IND	SKA	AND	C4	C2	C1	C3	C5	C6	C7	C8	C9	C10	C11	C12	C13
LSS	**0.16****	0.64*	0.82**	0.59×	0.29·	0.47·	0.15·	0.19·	0.51·	0.43·	0.36·	0.71·	0.64·	0.50·	−0.45·	−0.42·	0.53·
IND	0.34***	**0.52*****	0.59*	0.28·	0.84·	0.33·	−0.31·	0.18·	0.61×	0.82·	0.41·	0.56·	0.60·	0.69·	−0.90·	−0.52·	0.19·
SKA	0.36***	0.60***	**0.53*****	0.55×	0.36·	0.74·	0.14·	0.51·	0.54·	0.18·	0.28·	0.44·	0.74·	0.41·	−0.55·	−0.23·	0.57·
AND	0.25***	0.33***	0.45***	**0.44****	−0.07·	0.14·	0.49·	0.02·	0.54·	0.14·	−0.30·	0.48·	0.57·	0.21·	−0.31·	−0.16·	0.36·
C4	0.18**	0.73***	0.31***	−0.03·	**0.33**·	0.44·	−0.66·	0.20·	0.30·	0.60·	0.28·	0.34·	0.25·	0.47·	−0.78·	−0.65·	−0.17·
C2	0.25***	0.35***	0.67***	0.13·	0.29**	**0.50***	−0.26·	0.65·	0.12·	−0.14·	0.10·	0.20·	0.24·	−0.01·	−0.37·	−0.38·	0.15·
C1	0.02·	−0.10·	0.16×	0.65***	−0.35***	−0.06·	**0.54***	0.23·	0.47·	−0.39·	0.09·	−0.26·	0.50·	0.21·	0.28·	0.48·	0.76·
C3	0.22**	0.32**	0.53***	0.16·	0.23*	0.51***	0.21×	**0.47***	0.39·	−0.31·	0.40·	−0.33·	0.45·	0.19·	−0.35·	0.06·	0.56·
C5	0.08·	0.35***	0.50***	0.22*	0.22×	0.30**	0.30**	0.33***	**0.64****	0.29·	0.36·	0.06·	0.97·	0.85·	−0.58·	−0.01·	0.74×
C6	0.09	0.51***	0.19×	0.23*	0.25*	−0.13·	0.00·	−0.01·	0.28*	**0.46**·	0.28·	0.66·	0.25·	0.46·	−0.70·	−0.36·	−0.10·
C7	0.02·	0.34**	0.19×	0.18·	0.15·	−0.12·	0.24*	0.26*	0.16·	0.44***	**0.51**×	0.05·	0.42·	0.60·	−0.24·	−0.10·	0.58·
C8	0.20**	0.43***	0.25*	0.19×	0.27*	0.06·	−0.20×	0.00·	0.11·	0.58***	0.13·	**0.51***	0.13·	0.12·	−0.38·	−0.62·	−0.10·
C9	0.25**	0.39***	0.43***	0.40***	0.16·	0.21×	0.29*	0.40***	0.55***	0.26*	0.27*	0.16·	**0.56***	0.80·	−0.55·	0.00·	0.82·
C10	0.00·	0.21×	0.25*	0.14·	0.06·	0.08·	0.36**	0.27*	0.66***	0.33**	0.24*	−0.04·	0.36**	**0.42**·	−0.48·	−0.24·	0.61·
C11	−0.24**	−0.64***	−0.39***	−0.30**	−0.45***	−0.28*	0.03·	−0.35***	−0.20×	−0.30*	−0.19·	−0.22×	−0.41***	−0.12·	**0.33**·	0.32·	−0.13·
C12	−0.12·	−0.26*	−0.19×	−0.04·	−0.26*	−0.27*	0.29*	0.01·	0.15·	0.04·	−0.04·	−0.16·	0.04·	0.15·	0.23×	**0.39**·	0.25·
C13	0.04·	0.10·	0.35***	0.30**	−0.16·	0.10·	0.67***	0.47***	0.65***	0.18·	0.42***	−0.05·	0.51***	0.65***	−0.06·	0.36**	**0.66****

Following rules of signification were applied: × $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.
Abbreviations: LSS: Lab sensory score; IND, SKA and AND are compounds measured by ultra-high-performance liquid-chromatography tandem mass spectrometry; IND: Indole; SKA: Skatole; AND: Androstenone; C1-C13 are compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry; C1: Androstenone; C2: Skatole; C3: 2-Aminoacetophenone; C4: Indole; C5: Benzonitrile; C6: 2,5-Dimethyl-pyrazine; C7: 1-Methoxy-4-methyl-benzene; C8: 2-Propenyl-benzene; C9: 2,3-dimethylphenyl isocyanate; C10: Benzenemethanol; C11: Biphenyl; C12: Pyrazine; C13: 2-Dodecanone.

stayed non-significant (5). Genetic correlations of VOC traits were non-significant or tend to significance ($p < 0.1$). To the author's knowledge, no previous studies have reported any genetic correlations for most of these compounds. Although the stable execution of the REML algorithm (Fig. 2) during the estimation of covariance, we currently cannot make reliable assumptions based on the results.

3.3. Establishing selection index equations

The transformation matrix (**T**) obtained in this process and applied to 17 traits (i.e., LSS, ISA and VOC traits) to generate the pooled boar taint traits (i.e., ISA index and VOC index) represented the following relative contributions to LSS as reported in Table 3. This approach enables the calculation of coefficients using the covariance vector between the trait of interest (LSS) and the other traits (ISA or VOCs), together with the covariance matrix among the explanatory traits (ISA or VOCs). It allows the partial effect of each explanatory variable on LSS to be obtained while taking into account the interdependencies among these variables. These coefficients reflect the influence of each ISA or VOC compound on LSS, depending on the index, and allow the isolation of the direct effect of each variable. For each index, the coefficients were related to the total prediction of LSS.

The relative contributions of ISA traits to LSS were as follows: SKA contributed 34.3 %, IND 21.8 %, and AND being slightly less important (17.9 %), resulting in a total of 74.0 % all combined. These results were supported by previous studies (e.g., Bonneau et al., 2000; Burgeon et al., 2023a; Mathur et al., 2012; Markey et al., 2025; Mörlein et al., 2016a). These studies also highlighted also that SKA is a higher contributor to the boar taint phenotype than AND. These coefficients in this represented the relative importance of the relation between ISA or VOC traits and LSS and use the genetic relationships between these traits to build the index. This is why, although the two traits may be correlated, the contribution from partial coefficients may differ greatly. Therefore, these coefficients more accurately reflect the relationships between traits because they take into account the structural dependence between the compounds.

Table 3
Relative contributions to LSS of each compound for the two indices.

	Compounds	Relative contribution to LSS
UHPLC-MS/MS	IND	21.8 %
	SKA	34.3 %
	AND	17.9 %
	Total	74.0 %
HS-SPME-GCxGC-TOFMS	C1	3.6 %
	C2	14.4 %
	C3	1.6 %
	C4	2.2 %
	C5	2.9 %
	C6	6.2 %
	C7	2.9 %
	C8	24.1 %
	C9	9.1 %
	C10	3.9 %
	C11	0.8 %
	C12	6.3 %
	C13	12.3 %
	Total	90.3 %

Abbreviations: LSS: Lab sensory score; IND, SKA and AND are compounds measured by ultra-high-performance liquid-chromatography tandem mass spectrometry (UHPLC-MS/MS); IND: Indole; SKA: Skatole; AND: Androstenone; C1-C13 are compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry (HS-SPME-GCxGC-TOFMS); C1: Androstenone; C2: Skatole; C3: 2-Aminoacetophenone; C4: Indole; C5: Benzonitrile; C6: 2,5-Dimethyl-pyrazine; C7: 1-Methoxy-4-methyl-benzene; C8: 2-Propenyl-benzene; C9: 2,3-dimethylphenyl isocyanate; C10: Benzenemethanol; C11: Biphenyl; C12: Pyrazine; C13: 2-Dodecanone.

On the other hand, the relative contributions of VOC traits to LSS vector were ranging from over 20 % for C8 to below 1 % for C11, for a total of 90.3 % when considering all VOCs together. Most important compounds were C8, C2 (i.e., SKA equivalent) and C13 (>10 %), but many other compounds contributed between 2 % and 9 %. Even if these results are preliminary, when compared to the other 16 traits, C8 had a strong genetic contribution to LSS. As previously indicated C2, as its SKA equivalent, is also one of the main contributors which highlighted again this compound for its crucial importance.

The observed differences in contributions between ISA index and VOC index can be explained by various factors such as initial measurement methods and interdependencies between compounds. The ISA index includes well-known indolic relationships, i.e., SKA and IND, as well as AND which are interdependent into synthesis systems. However, when other compounds are present, the coefficients indicate only the relative importance within the index, taking into account the differences in compounds volatilisation and their interactions, as well as the interdependence of correlations between compounds. For example, C4 and C3 (i.e., IND equivalent and 2-aminoacetophenone) participate in the synthesis and degradation cycle of C2 (i.e., SKA equivalent) (Wesoly and Weiler, 2012) and C2 is also influenced by C1 (i.e., AND equivalent) regulation, including at the level of gene expression (Zamaratskaia and Squires, 2009.; Duarte et al., 2021). This implies that these compounds are genetically interrelated and contribute to the VOC index. For C8, but also other VOCs, these findings could be due to metabolic pathways, sensory interaction (e.g., C8 or similar aromatic compounds might influence the sensory perception of boar taint by our assessor) or even joint genetic x environmental factors (e.g., food affecting jointly the levels of aromatic compounds in pigs, which in turn could correlate genetics with the levels of boar taint compounds). Obviously, these contributions are very preliminary results, but above all they indicate that all VOC traits together contribute to LSS, and thus boar taint perception, with a variable relative importances.

3.4. Estimating genetic parameters using ISA and VOC indices

Heritabilities, genetic and phenotypic correlations of LSS, ISA index and VOC index traits are reported in Table 4. The heritabilities were moderate to high (h^2_{LSS}= 0.15±0.04; $h^2_{ISA\ index}$= 0.52±0.12; $h^2_{VOC\ index}$= 0.61±0.13). Genetic correlations were high (0.86–0.97) and phenotypic correlations were moderate to high (0.30–0.72). Their p-values were highly significant for all heritabilities, genetic and phenotypic correlations. It must be stressed that p-values were strongly reduced when comparing the estimates relative to the VOC index with the initial (co)variance components relative to VOC traits which leads to high significance of relationships (see Table 4 vs. Table 2). This supports the hypothesis that considering all the VOC traits together, i.e., pooling them, enhances their usefulness as a novel predictor of LSS. Although the small dataset introduces considerable uncertainty in the genetic correlation estimates, pooling them partially reduces this uncertainty. The usual traits, LSS and ISA, were highly genetically correlated as reported previously for the separate traits (Duarte et al., 2021; Markey et al., 2025). Although selection based on LSS as a single trait would be less accurate due to its lower heritability, this phenotype is widely available, less time-consuming and takes into account synergetic effects between compounds (Burgeon et al., 2021a). It reflects at least partially the “boar taint” in consumers’ perceptions (Mörlein et al., 2014). Haberland et al. (2014), found to be useful for genomic selection against boar taint. Comparing the VOC index to the ISA index, genetic correlations with LSS were higher for VOC index. As also heritability tended to be higher for the VOC index, considering only this aspect, the use of VOC for genetic selection could be preferred to ISA. The relative heritability of the VOC index indicated a predominant genetic influence on their presence. A strong genetic correlation between LSS with the VOC index indicated their strong influence on LSS. Using standard SI-formulas (Van Vleck, 1993) to estimate the correlated response on LSS, the relative advantage of also selecting for the VOC index can be approximated as 1.2 times that of using also the ISA index. This is based on a correlated response of 192 % for VOC index compared to 160 % for ISA index, both expressed relative to the direct response of LSS here put to 100 %. Moreover, strong genetic relationships appear between the ISA and VOC indices, as evidenced by their high genetic correlations, which are partially explained by their common composition. However, other influences of some other VOCs cannot be excluded. Considering odour in its globality and the knowledge already acquired on the subject of these compounds, boar taint is characterized by faecal (SKA, IND) and urinary (AND) flavour, but also foxy (2-aminoacetophenone) (Mörlein et al., 2024), roasted (C6) (Gerlach et al., 2016), floral (C10) and ‘orange, grass, fresh’ odours (C13) (Rodrigues et al., 2024). Indeed, the VOC olfactory activity, their volatility, interactions and concentrations should be studied for their influence on perception too (Mörlein et al., 2014, 2016a, 2024). Such in-depth studies are now possible thanks to technological progress (Burgeon, 2022).

Table 4
Heritabilities (on the diagonal), genetic (above) and phenotypic (below) correlations of lab sensory score, indole, skatole and androstenone measured by ultra-high-performance liquid-chromatography tandem mass spectrometry based, and compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry based indices.

	LSS	ISA index	VOC index
LSS	**0.15*****	0.86***	0.95***
ISA index	0.40***	**0.52*****	0.97***
VOC index	0.30***	0.72***	**0.61*****

Following rules of signification were applied: × $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Abbreviations: LSS: Lab sensory score; ISA: Indole, skatole and androstenone measured by ultra-high-performance liquid-chromatography tandem mass spectrometry; VOC: Volatile organic compounds measured by headspace solid-phase microextraction in combination with comprehensive two-dimensional gas chromatography–time–of–flight mass spectrometry.

3.5. Limitations and implications for future studies

Our exploratory study produced preliminary results that, although encouraging, require the availability of more data to be confirmed and expanded. We believe that, even if the amount of data is particularly limited, the publication of these preliminary results can generate further interest and contribute to existing knowledge. Moreover, the research presenter here goes beyond most previous studies which only focuses on analysis and phenotyping, by exploring the underlying genetic structure associated with the genetic expression of the complex phenotype represented by compounds related to perception of boar tainted pig fat. However, the data presented here suggests that this phenotype is more complex than it seems. This complexity is not yet fully integrated into current statistical approaches either. This finding highlights the need for methodological tools that can account for this multidimensionality, as we did by developing pooling approaches. These prospects highlight the importance of pursuing further research with larger and more diverse samples to validate the present findings on the usefulness of pooled phenotypes. Several directives for future work can be identified to expand and refine our study. A first improvement would be increasing the number of VOC records to increase reliability of the variance component estimations associated with these traits. Otherwise, increasing the number of assessors for sensory evaluations should improve knowledge on associated perception with compounds. Indeed, the boar taint characterization by assessors differs for each person. The presence of other potential compounds responsible for boar taint perception or amplifying its perception then studied, was reported in VOC profiles studies (Burgeon et al., 2021b, 2023a, 2023b; Rius et al.,

2005; Rius Solé and Garcia Regueiro, 2001). This also highlights the limitations of sensory analyses, which depend on protocols that influence perception and VOC release, accuracy and reliability of the method (Heyrman et al., 2020; Mörlein et al., 2014; Mathur et al., 2012). As previously demonstrated by different studies (e.g., Haberland et al., 2014; Markey et al., 2025; Windig et al., 2012), the use of sensory scores in a genetic context is feasible, and multidimensionality assessed by several different evaluators is a potential solution to this issue. As shown in Markey et al. (2025), optimizing the combination of assessors by using three best combination of assessors out of ten, enabled the generation of a qualitative boar taint reference phenotype with an heritability up to 0.40. Additionally, other largely unexplored phenotyping tools exist such as electronic noses and sensors (Burgeon et al., 2021a). Those tools could take into account the multidimensionality of the boar taint phenotype. Currently however, most of these novel tools have not been the investigated in the context of genetic studies to determine their usefulness in selection. These tools could help to overcome the major issue to obtain enough data to calculate more robust individual genetic parameters. More data would also allow to do genome wide association studies (Botelho et al., 2022; Groβe-Brinkhaus et al., 2015; Faggion et al., 2023) which could on one hand allow complementary research on the genetic background of these compounds and on the other hand develop efficient gene-informed genomic selection (Zheng et al., 2024). Finally, an area in which future research could future refine our study is the pooling strategy. Although our coefficients were based on an optimised genetic prediction of LSS, the reference trait remained phenotypic. Consequently, the pooling strategy cannot be easily applied to heterogeneous datasets (such as is the case here with missing data for some traits to assess the boar taint phenotype). Yet, this situation frequently arises in practical settings, as it is the case here in routine selection against boar taint, that often excludes costly analysis to be performed on all animals, leading to heterogeneous dataset.

4. Conclusion

This exploratory study generated preliminary, yet insightful results revealing that the previously identified novel compounds could contribute to the genetic evaluation of boar taint. This study confirmed that ISA traits are moderate to high heritable and strongly correlated with LSS. In comparison, new VOCs were highlighted for their potential genetic relationship with boar taint (i.e., LSS and/or ISA) and at least similar level of heritability as ISA traits. However, given the limited data availability in our exploratory study, the estimation of variance components, while stable and convergent, lead to non-significant relations for certain traits. The use of a SI-based pooling method allowed to estimate variance components for indices pooling ISA and VOC traits and to study their response to LSS, chosen as reference. Some remarkable contribution (>10 %) from compounds as C2 (equivalent SKA) but not exclusively, were used in the prediction equation. After estimation of components for indices, the results showed that the selection on VOC index could generate a strong response for LSS (i.e., reducing boar taint risk), stronger than when using the ISA index. These results highlight the potential use of VOCs, currently obtained by relatively laborious and costly methods. However, technologies are progressing and, under the condition such as low price, high-throughput could be developed, pooled VOCs could replace ISA traits in routine breeding applications. Moreover, the results demonstrated that this indirect response to selection would be stronger using VOC index than with the ISA index. More research is required, but this exploratory study already suggested that there is reasonable evidence to collect more VOC data enabling further research.

Data and model availability statement

The datasets and model are available by contacting authors on reasonable request and with permission of the breeding organizations.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors did not use generative AI or AI-assisted technologies in the writing process.

Funding

This research was supported by the Public Service of Wallonia (SPW) Agriculture of the Walloon Region (RW), Belgium, through grants D65–1430 and D31–1396 (NoWallOdor project), the Walloon Pig Genomic Evaluation System and the Fund for Scientific Research (FNRS), Belgium, through the grants PDR T.1053.15 (GENODOMICS project) providing genotypes and PDR T.0095.19 (DEEPSELECT project) providing computational resources. Author Anaïs Rodrigues was supported by the University of Liege under Special Funds for Research, IPD-STEMA Programme.

CRediT authorship contribution statement

Alice Markey: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing. **Anaïs Rodrigues:** Data curation, Formal analysis, Funding acquisition, Investigation, Writing – review & editing. **Pierre-Hugues Stefanuto:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Jean-François Focant:** Data curation, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – review & editing. **Anne-Catherine Huet:** Funding acquisition, Project administration, Resources, Writing – review & editing. **José Wavreille:** Funding acquisition, Project administration, Resources, Writing – review & editing. **Katrien Wijnrocx:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Nicolas Gengler:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Nicolas Gengler reports financial support was provided by Public Service of Wallonia (SPW) Agriculture of the Walloon Region (RW). Nicolas Gengler reports financial support was provided by Fund for Scientific Research (FNRS). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank all contributors in the NoWallOdor project. We especially acknowledge the important contribution of Elevéo (Ciney, Belgium) in the breeding, raising and recording of animals.

References

Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., Lawlor, T.J., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy. Sci. 93, 743–752. https://doi.org/10.3168/jds.2009-2730.

Aguilar, I., Misztal, I., Legarra, A., Tsuruta, S., 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J. Anim. Breed. Genet. 128, 422–428. https://doi.org/10.1111/j.1439-0388.2010.00912.x.

Bejerholm, C., Barton-Gade, P., 1993. The relationship between skatole/androstenone and odour/flavour of meat from entire male pigs. In measurement and prevention of

boar taint in entire male pigs. In: Bonneau, M. (Ed.), Proc. EAAP Working Group: Production and Utilisation of Meat from Entire Male Pigs. INRA, Paris, pp. 75–79.
Berg, H., Agerhem, H., von Seth, G., Tornberg, E., Andresen, Ø., 1993. The relationship between skatole and androstenone content and sensory off-odour in entire male pigs. In measurement and prevention of boar taint in entire male pigs. In: Bonneau, M. (Ed.), Proc. EAAP Working Group: Production and Utilisation of Meat from Entire Male Pigs. INRA, Paris, pp. 55–61.
Botelho, M.E., Lopes, M.S., Mathur, P.K., Knol, E.F., Fonseca e Silva, F., Lopes, P.S., Gimarães Marques, D.B.D., Veroneze, R., 2022. Weighted genome-wide association study reveals new candidate genes related to boar taint compounds. Livest. Sci. 257, 104845. https://doi.org/10.1016/j.livsci.2022.104845.
Bonneau, M., 1993. Effects of different compounds on boar taint. In: 44th Annual Meeting of the EAAP. Denmark, pp. 326–327.
Bonneau, M., Walstra, P., Claudi-Magnussen, C., Kempster, A.J., Tornberg, E., Fischer, K., Diestre, A., Siret, F., Chevillon, P., Claus, R., Dijksterhuis, G., Punter, P., Matthews, K.R., Agerhem, H., Béague, M.P., Oliver, M.A., Gispert, M., Weiler, U., von Seth, G., Leask, H., Font-i-Furnols, M., Homer, D.B., Cook, G.L., 2000. An international study on the importance of androstenone and skatole for boar taint: IV. Simulation studies on consumer dissatisfaction with entire male pork and the effect of sorting carcasses on the slaughter line, main conclusions and recommendations. Meat. Sci. 54 (3), 285–295. https://doi.org/10.1016/S0309-1740(99)00105-9.
Burgeon, C., 2022. Headspace Sampling of Boar Taint Compounds in Backfat for Detection Based on Volatile Organic Compounds. ORBi [Doctoral thesis, ULiège. GxABT]. https://orbi.uliege.be/handle/2268/294000.
Burgeon, C., Debliquy, M., Lahem, D., Rodriguez, J., Ly, A., Fauconnier, M.-L., 2021. Past, present, and future trends in boar taint detection. Trends. Food Sci. Technol. 112, 283–297. https://doi.org/10.1016/j.tifs.2021.04.007.
Burgeon, C., Font-i-Furnols, M., Garrido, M.D., Linares, M.B., Brostaux, Y., Sabeña, G., Fauconnier, M.-L., Panella-Riera, N., 2023ab. Can sensory boar taint levels be explained by fatty acid composition and emitted volatile organic compounds in addition to androstenone and skatole content? Meat. Sci. 195, 108985. https://doi.org/10.1016/j.meatsci.2022.108985.
Burgeon, C., Markey, A., Brostaux, Y., Fauconnier, M.-L., 2023b. Modifying headspace sampling environment improves detection of boar taint compounds in pork fat samples. Chemosensors 11, 551. https://doi.org/10.3390/chemosensors11110551.
Burgeon, C., Markey, A., Debliquy, M., Lahem, D., Rodriguez, J., Ly, A., Fauconnier, M.-L., 2021b. Comprehensive SPME-GC–MS analysis of VOC profiles obtained following high-temperature heating of pork back fat with varying boar taint intensities. Foods. 10, 1311. https://doi.org/10.3390/foods10061311.
Christensen, O.F., Lund, M.S., 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42, 2. https://doi.org/10.1186/1297-9686-42-2.
Duarte, D., Schroyen, M., Mota, R.R., Vanderick, S., Gengler, N., 2021. Recent genetic advances on boar taint reduction as an alternative to castration: a review. J. Appl. Genet. 62, 137–150. https://doi.org/10.1007/s13353-020-00598-w.
Eaton, J.W., Bateman, D., Hauberg, S., Wehbring, R., 2020. GNU Octave version 6.1.0 manual: a high-level interactive language for numerical computations. Available from. https://www.gnu.org/software/octave/doc/v6.3.0/.
Faggion, S., Boschi, E., Veroneze, R., Carnier, P., Bonfatti, V., 2023. Genomic prediction and genome-wide association study for boar taint compounds. Animals 13 (15), 2450. https://doi.org/10.3390/ani13152450.
Fischer, J., Elsinghorst, P.W., Bücking, M., Tholen, E., Petersen, B., Wüst, M., 2011. Development of a candidate reference method for the simultaneous quantitation of the boar taint compounds androstenone, 3α-androstenol, 3β-androstenol, skatole, and indole in pig fat by means of stable isotope dilution analysis–headspace solid-phase microextraction–gas chromatography/mass spectrometry. Anal. Chem. 83, 6785–6791. https://doi.org/10.1021/ac201465q.
Fischer, J., Gerlach, C., Meier-Dinkel, L., Elsinghorst, P.W., Boeker, P., Schmarr, H.G., Wüst, M., 2014. 2-Aminoacetophenone—A hepatic skatole metabolite as a potential contributor to boar taint. Food Res. Int. 62, 35–42. https://doi.org/10.1016/j.foodres.2014.02.045.
Gerlach, C., Elsinghorst, P.W., Schmarr, H.-G., Wüst, M., 2016. 2-Aminoacetophenone is the main volatile phase I skatole metabolite in Pietrain × Baden-Württemberg hybrid type boars. J. Agric. Food Chem. 61 (5), 1158–1163. https://doi.org/10.1021/acs.jafc.5b05729.
Gerlach, C., Leppert, J., Santiuste, A.C., Pfeiffer, A., Boeker, P., Wüst, M., 2018. Comparative aroma extract dilution analysis (cAEDA) of fat from tainted boars, castrated male pigs, and female pigs. J. Agric. Food Chem. 66 (10), 2403–2409. https://doi.org/10.1021/acs.jafc.6b04747.
Groβe-Brinkhaus, C., Storck, L.C., Frieden, L., Neuhoff, C., Schellander, K., Looft, C., Tholen, E., 2015. Genome-wide association analyses for boar taint components and testicular traits revealed regions having pleiotropic effects. BMC. Genet. 16, 36. https://doi.org/10.1186/s12863-015-0194-z.
Haberland, A.M., Luther, H., Hofer, A., Tholen, E., Simianer, H., Lind, B., Baes, C., 2014. Efficiency of different selection strategies against boar taint in pigs. Animal. 8, 11–19. https://doi.org/10.1017/S1751731113001857.
Hansen-Møller, J., 1994. Rapid high-performance liquid chromatographic method for simultaneous determination of androstenone, skatole and indole in back fat from pigs. J. Chromatogr. B: Biomed. Sci. Appl. 661 (2), 219–230. https://doi.org/10.1016/S0378-4347(94)80049-9.
Hansson, K.-E., Lundström, K., Fjelkner-Modig, S., Persson, J., 1980. The importance of androstenone and skatole for boar taint. Swed. J. Agric. Res. 10, 167–173.
Heyrman, E., Janssens, S., Buys, N., Vanhaecke, L., Millet, S., Tuyttens, F.A.M., Wauters, J., Aluwé, M., 2020. Developing and understanding olfactory evaluation of boar taint. Animals 10 (9), 1684. https://doi.org/10.3390/ani10091684.
Lin, C.Y., 1978a. Estimation of heritability of an index. Can. J. Genet. Cytol. 20, 485–487. https://doi.org/10.1139/g78-056.
Lin, C.Y., 1978b. Index selection for genetic improvement of quantitative characters. Theor. Appl. Genet. 52, 49–56. https://doi.org/10.1007/BF00281316.
Makridis, G., Heyrman, E., Kotios, D., Mavrepis, P., Callens, B., Van De Dijver, R., Maselyne, J., Aluwé, M., Kyriazis, D., 2022. Evaluating machine learning techniques to define the factors related to boar taint. Livest. Sci. 264, 105045. https://doi.org/10.1016/j.livsci.2022.105045.
Markey, A., Groβe-Brinkhaus, C., Mörlein, D., Mörlein, J., Wilmot, H., Tholen, E., Gengler, N., 2025. Genetic investigations into the use of sensory evaluation: the case of boar taint discrimination in Pietrain sired crossbreds. J. Anim. Sci. 103. https://doi.org/10.1093/jas/skae389.
Markey, A., Wavreille, J., Mayeres, P., Huet, A.C., Gengler, N., 2023. PSVI-5 optimizing limited deep phenotyping capacity for genetic studies and routine genetic evaluations: application to boar taint. J. Anim. Sci. 101 (S3), 406–407. https://doi.org/10.1093/jas/skad281.482.
Mathur, P.K., ten Napel, J., Bloemhof, S., Heres, L., Knol, E.F., Mulder, H.A., 2012. A human nose scoring system for boar taint and its relationship with androstenone and skatole. Meat. Sci. 91, 414–422. https://doi.org/10.1016/j.meatsci.2012.02.025.
Meyer, K., Houle, D., 2013. Sampling based approximation of confidence intervals for functions of genetic covariance matrices. In: 20th Association Advancement Animal Breeding Genetics Conference 20-23 October 2013. Napier, New Zealand, pp. 523–526.
Meyer, K., Kirkpatrick, M., 2010. Better estimates of genetic covariance matrices by "bending" using penalized maximum likelihood. Genetics 185, 1097–1110. https://doi.org/10.1534/genetics.109.113381.
Misztal, I., 2008. Reliable computing in estimation of variance components. J. Anim. Breed. Genet. 125, 363–370. https://doi.org/10.1111/j.1439-0388.2008.00774.x.
Misztal, I., Tsuruta, S., Lourenco, D.A.L., Masuda, Y., Aguilar, I., Legarra, A., Vitezica, Z., 2014. Manual for BLUPF90 Family Programs. University of Georgia. Retrieved on 21/10/2024 from. https://nce.ads.uga.edu/wiki/doku.php?id=documentation.
Mörlein, D., Mörlein, J., Gerlach, C., Strack, M., Kranz, B., Brüggemann, D.A., 2024. An overlooked compound contributing to boar taint and consumer rejection of meat products: 2-aminoacetophenone. Meat. Sci. 213, 109497. https://doi.org/10.1016/j.meatsci.2024.109497.
Mörlein, D., Mörlein, J., Gertheiss, J., Meier-Dinkel, L., Fischer, J., Eynck, H.-J., Heres, L., Looft, C., Tholen, E., 2016a. Interaction of skatole and androstenone in the olfactory perception of boar taint. J. Agric. Food Chem. 64 (22), 4556–4565. https://doi.org/10.1021/acs.jafc.6b00355.
Mörlein, J., Gertheiss, J., Wicke, M., Mörlein, D., 2014. How olfactory acuity affects the sensory assessment of boar fat: a proposal for quantification. Meat. Sci. 98, 255–262. https://doi.org/10.1016/j.meatsci.2014.05.037.
Mörlein, J., Meier-Dinkel, L., Gertheiss, J., Mörlein, D., 2016b. Boar taint detection: a comparison of three sensory protocols. Meat. Sci. 111, 92–100. https://doi.org/10.1016/j.meatsci.2015.08.011.
Mujibi, F.D.N., Crews, D.H., 2009. Genetic parameters for calving ease, gestation length, and birth weight in Charolais cattle. J. Anim. Sci. 87, 2759–2766. https://doi.org/10.2527/jas.2008-1141.
Rius, M.A., Hortós, M., García-Regueiro, J.A., 2005. Influence of volatile compounds on the development of off-flavours in pig back fat samples classified with boar taint by a test panel. Meat. Sci. 71, 595–602. https://doi.org/10.1016/j.meatsci.2005.03.014.
Rius Solé, M.A., García Regueiro, J.A., 2001. Role of 4-phenyl-3-buten-2-one in boar taint: identification of new compounds related to sensorial descriptors in pig fat. J. Agric. Food Chem. 49, 5303–5309. https://doi.org/10.1021/jf010482d.
Rodrigues, A., Massenet, T., Dubois, L.M., Huet, A.-C., Markey, A., Wavreille, J., Gengler, N., Stefanuto, P.-H., Focant, J.-F., 2024. Development and validation of a classification model for boar taint detection in pork fat samples. Food Chem. 443, 138572. https://doi.org/10.1016/j.foodchem.2024.138572.
SAS Institute Inc., 2025. SAS 9.4. SAS Institute, Cary, NC, USA.
Sørensen, K.M., Westley, C., Goodacre, R., Engelsen, S.B., 2015. Simultaneous quantification of the boar-taint compounds skatole and androstenone by surface-enhanced Raman scattering (SERS) and multivariate data analysis. Anal. Bioanal. Chem. 407 (25), 7787–7795. https://doi.org/10.1007/s00216-015-8945-2.
Snell, E.J., 1964. A scaling procedure for ordered categorical data. Biometrics 20, 592–607. https://doi.org/10.2307/2528498.
Squires, E.J., Bone, C., Cameron, J., 2020. Pork production with entire males: directions for control of boar taint. Animals 10 (9), 1665. https://doi.org/10.3390/ani10091665.
Tajet, H., Andresen, O., Meuwissen, T., 2006. Estimation of genetic parameters of boar taint; skatole and androstenone and their correlations with sexual maturation. Acta Vet. Scand. 48 (S9). https://doi.org/10.1186/1751-0147-48-S1-S9.
Van Vleck, L.D., 1993. Selection Index and Introduction to Mixed Model Methods for Genetic Improvement of animals: the Green Book. CRC Press, Boca Raton, FL, USA.
VanRaden, M.P., Van Tassell, P.C., Wiggans, C.G., Sonstegard, S.T., Schnabel, D.R., Taylor, F.J., Schenkel, S.F., 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. J. Dairy. Sci. 92, 16–24. https://doi.org/10.3168/jds.2008-1514.
Verplanken, K., Wauters, J., Van Durme, J., Claus, D., Vercammen, J., De Saeger, S., Vanhaecke, L., 2016. Rapid method for the simultaneous detection of boar taint

compounds by means of solid phase microextraction coupled to gas chromatography/mass spectrometry. J. Chromatogr. A 1462, 124–133. https://doi.org/10.1016/j.chroma.2016.07.077.

Wang, H., Misztal, I., Aguilar, I., Legarra, A., Muir, W.M., 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. (Camb) 94, 73–83. https://doi.org/10.1017/S0016672312000274.

Wesoly, R., Weiler, U., 2012. Nutritional influences on skatole formation and skatole metabolism in the pig. Animals 2 (2), 221–242. https://doi.org/10.3390/ani2020221.

Windig, J.J., Mulder, H.A., ten Napel, J., Knol, E.F., Mathur, P.K., Crump, R.E., 2012. Genetic parameters for androstenone, skatole, indole, and human nose scores as measures of boar taint and their relationship with finishing traits. J. Anim. Sci. 90, 2120–2129. https://doi.org/10.2527/jas.2011-4700.

Zamaratskaia, G., Squires, E.J., 2009. Biochemical, nutritional and genetic effects on boar taint in entire male pigs. Animal. 3, 1508–1521. https://doi.org/10.1017/S1751731108003674.

Zheng, Z., Liu, S., Sidorenko, J., Wang, Y., Lin, T., Yengo, L., Turley, P., Ani, A., Wang, R., Nolte, I.M., Snieder, H., Study, L.L.C., Yang, J., Wray, N.R., Goddard, M.E., Visscher, P.M., Zeng, J., 2024. Leveraging functional genomic annotations and genome coverage to improve polygenic prediction of complex traits within and between ancestries. Nat, Genet. 56, 767–777. https://doi.org/10.1038/s41588-024-01704-y.