PRIMJENA GENOMSKIH INFORMACIJA U SELEKCIJI NERASTA NA KONCENTRACIJE SKATOLA I ANDROSTENONA

Lukić, Boris

REPUBLIC OF CROATIA
J.J. STROSSMAYER UNIVERSITY OF
OSIJEK
FACULTY OF AGRICULTURE IN OSIJEK

Boris Lukić, dipl.ing.

IMPLEMENTATION OF GENOMIC INFORMATION IN SELECTION OF BOARS ON SKATOLE AND ANDROSTENONE CONCENTRATIONS

PhD THESIS

Osijek, 2015

REPUBLIC OF CROATIA
J.J. STROSSMAYER UNIVERSITY OF
OSIJEK
FACULTY OF AGRICULTURE IN OSIJEK

Boris Lukić, dipl.ing.

IMPLEMENTATION OF GENOMIC INFORMATION IN SELECTION OF BOARS ON SKATOLE AND ANDROSTENONE CONCENTRATIONS

- PhD thesis -

Osijek, 2015

REPUBLIC OF CROATIA
J.J. STROSSMAYER UNIVERSITY OF
OSIJEK
FACULTY OF AGRICULTURE IN OSIJEK

Boris Lukić, dipl.ing.

IMPLEMENTATION OF GENOMIC INFORMATION IN SELECTION OF BOARS ON SKATOLE AND ANDROSTENONE CONCENTRATIONS

- PhD thesis -

Mentor: prof. dr. sc. Goran Kušec

Evaluation committee:

1. **Dr. sc. Sonja Jovanovac, Full professor, Faculty of Agriculture, J.J. Strossmayer University of Osijek, chair**
2. **Dr.sc. Goran Kušec, Full professor, Faculty of Agriculture in Osijek, J.J. Strossmayer University of Osijek, mentor**
3. **Dr.sc. John Woolliams, Full professor, Roslin Institute, University of Edinburgh, co-mentor**
4. **Dr. sc. Ino Čurik, Full professor, Faculty of Agriculture, University of Zagreb, member**
5. **Dr. sc. Ivona Djurkin Kušec, Assistant professor, Faculty of Agriculture, J.J. Strossmayer University of Osijek, member**

Osijek, 2015

REPUBLIC OF CROATIA
J.J. STROSSMAYER UNIVERSITY OF
OSIJEK
FACULTY OF AGRICULTURE IN OSIJEK

Boris Lukić, dipl.ing.

IMPLEMENTATION OF GENOMIC INFORMATION IN SELECTION OF BOARS ON SKATOLE AND ANDROSTENONE CONCENTRATIONS

- PhD thesis -

Mentor: prof. dr. sc. Goran Kušec

Date of the thesis defence was assigned to be held on 8th of May, 2015 in front of the committee members:

1. **Dr. sc. Sonja Jovanovac, Full professor, Faculty of Agriculture, J.J. Strossmayer University of Osijek, chair**
2. **Dr.sc. Goran Kušec, Full professor, Faculty of Agriculture in Osijek, J.J. Strossmayer University of Osijek, mentor**
3. **Dr.sc. John Woolliams, Full professor, Roslin Institute, University of Edinburgh, co-mentor**
4. **Dr. sc. Ino Čurik, Full professor, Faculty of Agriculture, University of Zagreb, member**
5. **Dr. sc. Ivona Djurkin Kušec, Assistant professor, Faculty of Agriculture, J.J. Strossmayer University of Osijek, member**

Osijek, 2015

BASIC DOCUMENTATION CARD

J.J. Strossmayer University of Osijek

PhD thesis

Faculty of Agriculture in Osijek
Postgraduate doctoral study: AGRICULTURAL SCIENCES
Course: Animal breeding

UDK: 636.082.2:636.4
Scientific area: Biotechnical sciences
Scientific field: Agriculture
Major: Animal breeding

IMPLEMENTATION OF GENOMIC INFORMATION IN SELECTION OF BOARS ON SKATOLE AND ANDROSTENONE CONCENTRATIONS

Boris Lukić, dipl.ing.

This thesis is written on Faculty of Agriculture in Osijek, University of J.J. Strossmayer in Osijek, in collaboration with the Roslin Institute, University of Edinburgh
Mentor: prof. dr. sc. Goran Kušec

Genomic predictors offer an opportunity to overcome the limitations of classical selection against boar taint, and this study evaluated different approaches to obtain such predictors. Samples from 941 pigs were included in a design which was dominated by 421 sib pairs, each pair having an animal with a high and a low skatole concentration (≥0.3 μg/g). All samples were measured for skatole and androstenone and genotyped using the Illumina SNP 60K porcine Illumina beadchip. The accuracy of predicting phenotypes was assessed by cross-validation using six different genomic evaluation methods, GBLUP and five Bayesian methods. The range of accuracies obtained by different prediction methods was narrow for androstenone, between 0.29 (Bayes Lasso) and 0.31 (Bayes B), and wide for skatole, between 0.21 (GBLUP) and 0.26 (Bayes SSVS). Relative accuracies corrected for h^2, were 0.54-0.56 and 0.75-0.94 for androstenone and skatole, respectively. The whole genome evaluation methods gave greater accuracy than using QTL alone (one SNP for androstenone and one SNP for skatole). Also, the dominance genetic variation was ignored in national evaluations, so we estimated the dominance genetic variance for androstenone and skatole using SNP information. For androstenone, GBLUP with dominance effects included captured substantial ratio of the dominance genetic variances (13%) in total variation. For skatole, more dominance genetic variance was captured by regional chromosomal heritability approach, particularly on chromosome 9, where the proportion of chromosomal dominance genetic variance in total dominance variance was 96%. The results demonstrate that GBLUP for androstenone is the simplest genomic technology to implement and one of the most accurate methods while more specialised models may be preferable for skatole. Dominance genetic effects included could provide additional source of genetic variation for both traits, therefore it is worthwhile considering in genomic evaluations.

Number of pages: 110
Number of figures: 19
Number of tables: 13
Number of literature citations: 117
Language: English
Key words: boar taint, skatole, androstenone, genomic selection, GBLUP, Bayes, dominance, regional heritability
Date of the thesis defence:
Thesis committee:

1. **prof. dr. sc. Sonja Jovanovac** - chair
2. **prof. dr. sc. Goran Kušec** - mentor
3. **prof. dr. sc. John Woolliams –** co-mentor
4. **prof. dr. sc. Ino Čurik** - member
5. **doc. dr. sc. Ivona Djurkin Kušec** - member

Thesis deposited in: National and University Library in Zagreb; J.J. Strossmayer University of Osijek; City and University Library Osijek, Faculty of Agriculture (Library)

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište Josipa Jurja Strossmayera u Osijeku

Doktorski rad

Poljoprivredni fakultet u Osijeku
Poslijediplomski doktorski studij: Poljoprivredne znanosti
Smjer: Stočarstvo

UDK: 636.082.2:636.4
Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Poljoprivreda
Grana: Stočarstvo

PRIMJENA GENOMSKIH INFORMACIJA U SELEKCIJI NERASTA NA KONCENTRACIJE SKATOLA I ANDROSTENONA

Boris Lukić, dipl.inž.

Rad je izrađen na Poljoprivrednom fakultetu Sveučilišta Josipa Jurja Strossmayera u Osijeku te Roslin institutu Sveučilišta u Edinburgu
Mentor: prof. dr. sc. Goran Kušec

Genomski markeri odnosno prediktori, pružaju nove mogućnosti u prevladavanju dosadašnjih ograničenja klasične selekcije protiv nerastovskog svojstva stoga je cilj ovog istraživanja istražiti mogućnosti njihove primjene. U istraživanje je uključen 941 nerast, od kojih je braće po ocu i majci, odnosno uparenih srodnika 421. Nadalje, svaki nerast s visokom razinom skatola ima brata s niskom razinom skatola u masnom tkivu (≥0.3 ug/g). Svim uzorcima su osim razine skatola utvrđene i razine androstenona. Životinje su genotipizirane koristeći Illumina SNP60K porcine beadchip. Analiza unakrsne provjere je izvršena s ciljem uspoređivanja točnosti procjene GBLUP metode i pet Bayes metoda na osnovi regresije u procjenjivanju nepromatranih fenotipskih svojstava. Raspon ostvarenih točnosti koristeći različite metode procjena je bio uzak za androstenon, između 0,29 (Bayes Lasso) i 0,31 (Bayes B), te širi za skatol, između 0,21 (GBLUP) i 0,26 (Bayes SSVS). Relativne točnosti korigirane s prosječnim h^2, iznosile su 0,54 - 0,56 za androstenon i 0,75-0,94 za skatol. S obzirom na to da se učinci dominantnosti u ukupnoj genetskoj varijabilnosti uglavnom nisu promatrali u nacionalnim uzgojnim programima evaluacije, dodatni je pristup korišten za procjenu dominantne genetske varijance za androstenon i skatol. Za androstenon, GBLUP s uključenim dominantnim učincima u modelu procijenio značajne omjere dominantnih genetskih varijanci (13%) u ukupnoj varijabilnosti. U slučaju skatola, procijenjena dominantna genetska varijanca bila je viša pristupom regionalnog heritabiliteta, i to najviše na kromosomu 9, gdje je udio kromosomske regionalne dominantne varijance u ukupnoj dominantnoj varijanci iznosio 96%. Dobiveni rezultati pokazuju da je GBLUP najjednostavnija metoda genomske procjene za androstenon, ujedno je lako provediva te jedna od najtočnijih metoda. U slučaju skatola, potrebno je kreirati prilagođene modele koji mogu postići značajno bolje rezultate. Također, dominantni genetski učinci pružaju dodatne izvore genetske varijabilnosti za oba svojstva i istodobno nude novi i primjenjivi pristup u genomskim procjenama u okviru nacionalnih uzgojnih programa.

Broj stranica: 110
Broj slika: 19
Broj tablica: 13
Broj literaturnih navoda: 117
Jezik izvornika: engleski
Ključne riječi: nerastovsko svojstvo, skatol, androstenon, genomska selekcija, GBLUP, Bayes, dominantnost, regionalni heritabilitet
Datum obrane:.
Povjerenstvo za obranu:

1. **prof. dr. sc. Sonja Jovanovac** - predsjednica
2. **prof. dr. sc. Goran Kušec** - mentor
3. **prof. dr. sc. John Woolliams –** komentor
4. **prof. dr. sc. Ino Čurik** - član
5. **doc. dr. sc. Ivona Djurkin Kušec** – član

Rad je pohranjen u: Nacionalna i sveučilišna knjižnica u Zagrebu, Sveučilište Josipa Jurja Strossmayera u Osijeku, Gradska i sveučilišna knjižnica Osijek, Knjižnica Poljoprivrednog fakulteta u Osijeku

If you don't go after what you want, you'll never have it. If you don't ask, the answer is always no. If you don't step forward, you're always in the same place.

Nora Roberts

TABLE OF CONTENTS

1. INTRODUCTION

General background on boar taint

Boar taint is an unpleasant odour that emanates from boar fat when it is heated or cooked. Castration of male pigs has been a common practice for many years in pig production, mainly serving as a measure to avoid this undesirable trait (Squires and Schenkel, 2010). On the one hand, castration is effective in prevention of the occurrence of boar taint and it makes male animals calmer and less aggressive as well as their behaviour more easily controlled. On the other hand, castration of male animals negatively affects some important production traits such as average daily weight gain, feed conversion efficiency and meat percentage (Fowler et al., 1981; Bonneau and Squires, 2004). The lowering of growth rate and production efficiency can be explained by a lack of sex hormones such as testosterone due to the removal of the gonads. Furthermore, production costs of castrated male pigs are significantly higher, which reduces economic profit (de Lange and Squires 1995, Lin et al., 2006). In addition to these production drawbacks, the World Society for the Protection of Animals criticized surgical castration of animals in intensive livestock farming in the past decade, so in many countries castration had to be performed with anaesthetics or in less painful or less stressful ways.

Castration is still employed in most European countries as a simple and effective practice, but given the guidelines of the European Commission, this practice is changing. In some countries, such as Norway, castration is carried out with the use of analgesics as per legislation passed in 2002. In the Netherlands, the use of analgesics has been obligatory in extensive farming systems since 2007 (Fredriksen et al., 2009).

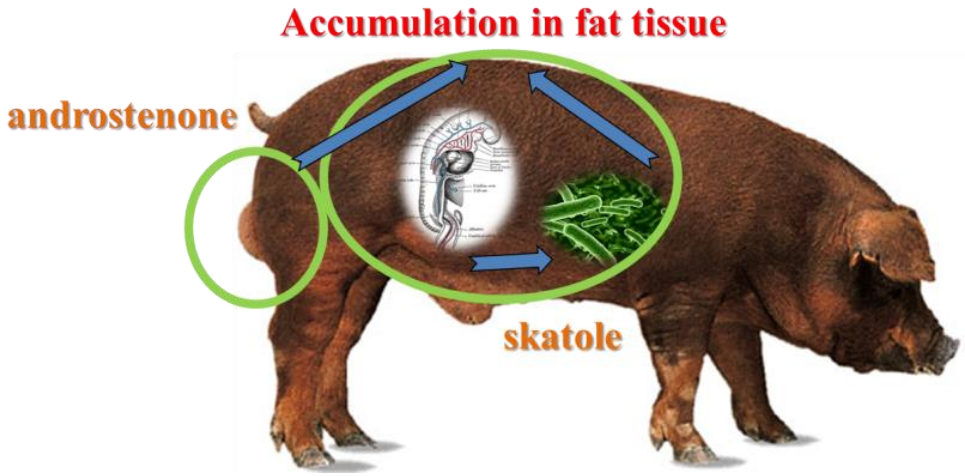
In Croatia, the Animal Protection Act (NN135/06, 37/13, 125/13) and the Ordinance laying down minimum standards for the protection of pigs (NN119/10) regulate the implementation of castration with regard to the problem of boar taint. The above regulations allow the castration of young boars, if carried out for zootechnical purposes by a veterinarian or a qualified person in accordance with the rules. Additionally, if the piglets are older than seven days, castration has to be performed by veterinarian using anaesthetics or analgesics.

According to the European Declaration on pig castration (European Declaration on alternatives to surgical castration of pigs, 2011), the first step in avoiding the practice of castration altogether began on the first of January 2012, requiring that castration should be performed with the use of analgesics or other painkillers. The second step requires that the practice of castrating young boars becomes completely abandoned from the first of January 2018 in all the member states of the European Union.

Causes of boar taint

Androstenone and skatole, which are considered responsible for the aforementioned unpleasant odour and taste of the meat, are chemical compounds that accumulate in fat tissue (Robic et al., 2008).

Androstenone is a steroid hormone produced in the testes at the beginning of sexual maturity, and is responsible for urine-like odour of meat. Skatole is a chemical substance produced from the amino acid tryptophan by work of bacteria in the colon and has a strong faecal odour. These two chemical compounds jointly contribute to an unpleasant smell and taste released during the cooking of meat (Babol et al., 1999; *Scheme 1*).



Scheme 1. Accumulation of androstenone and skatole

Possible alternatives

In the past 10 years, this issue has been extensively investigated with many published studies describing different approaches on the mitigation of boar taint. These studies focused on: identifying the responsible candidate genes and QTLs of these chemical components and their physiology; early detection with rapid analytical methods and sensory evaluations; immunocastration of piglets; animal nutrition; consumer acceptability studies and threshold levels of chemical substances in fat as well as the influence on the meat quality.

Several alternative approaches have been proposed for preventing boar taint (Bonneau and Squires, 2004). For example, immunocastration is one alternative involving vaccination to inhibit testicular function, but problems arise due to cost (de Roest et al., 2009), the need for repeated vaccination (Squires and Bonneau, 2004), variation in vaccine response (Bonneau et al., 1994 and Turkstra et al., 2002) and there are risks to male operatives from accidental self-inocculations. Other alternatives include slaughtering animals before sexual maturity, which is common practice in the UK, but is impractical in most EU countries for reasons of

consumer acceptability or profitability (Xue et al., 1997). A more acceptable and practical long-term approach is the genetic selection of animals against expression of boar taint (Quintanilla et al., 2003; Lee et al., 2005; Moe et al., 2009; Squires and Schenkel, 2010; Duijvesteijn et al., 2010; Rowe et al., 2014). Table 1 summarizes all the possible strategies of how to avoid tainted boar taint with regard to animal welfare, acceptability, risks and cost (EFSA Report, 2004).

Table 1. Strategies of how to avoid tainted boar taint with regard to animal welfare, acceptability, risks and cost (EFSA Report, 2004)

WELFARE	BOAR TAINT	PROCEDURES, CONDITIONS, ACCEPTBILITY AND RISKS	PRODUCTION, MEAT QUALITY	COST
General anaesthesia : risk for animals, discomfort at induction and awakening Local anaesthesia: transient pain at injection Handling stress	Effective against boar taint	General anaesthesia: no licensed anaesthetic, risk for people doing anaesthesia **Surgical castration with anaesthesia / analgesia** Public acceptance of invasive procedure and integrity of the animal	No production advantages of entire males (growth, leanness, efficiency) Easier management than for entire	Price of anaesthetics and analgesics Anaesthesia: cost of veterinary assistance
Chemical castration : pain and discomfort ? Immunocastration: minor pain at injections, discomfort?, hypothalamic lesions? Handling stress	Chemical castration : effectiveness to be ascertained Immunocastration: largely effective against boar taint, some animals do not respond	**Chemical castration** **Immunocastration** Minor risk of residues Risk of self-injection	Chemical castration: presumably advantages and drawbacks of castrates Early immunocastration: advantages and drawbacks of castrates Late immunocastration: production's advantages of entire males, management as for castrates?	Price of vaccine Workload of double vaccination Cost of screening /analysis of carcasses
No invasive procedures More fighting and sexual harassment	Unacceptable high frequency of tainted carcasses which must be reduced by various means	**Raising entire males** **Feeding :** to minimise skatole and avoid soft fat **Genetics:** choice of breed and selection for low taint **Management:** social environment, space, avoid mixing and competition, develop birth to slaughter system Cleanliness, facilities for thermoregulation	Poduction's advantages of entire males More complex management than for castrates Risk of carcass damage, inferior meat quality and soft fat Risk of taint Necessity to develop new products for use of tainted meat	Cost of analyses for skatole and androstenone on slaughter line Reduced value of tainted carcasses Cost of marking system if consumers are to have choice between meat from different types of carcasses

Implementation of genomic information in selection against boar taint

For these mentioned chemical components, androstenone and skatole reports on high values of heritability suggest the application of genetic selection against boar taint as method with considerable prospects. Also the positive genetic correlations (Tajet et al., 2006; Engelsma et al., 2007) between androstenone and skatole could make selection easier to implement. However, androstenone is a steroid hormone that shares physiological pathways with other steroids with which it has high correlations. Because of that, genetic selection resulted with lower levels of androstenone and reproductive problems in upcoming generations as a side effect. Previous attempts in genetic selection (Willeke et al., 1980; Willeke et al., 1987; Sellier et al., 1988; Willeke and Pirchner, 1989; Sellier et al., 2000) didn't show success mainly due to the low accuracy of estimated breeding values (EBVs).

Classical selection approach in animal breeding combines phenotypic measurements and probabilities (estimated from the pedigree data) that genes are identical by descent (Van Raden, 2009), and by this approach breeding schemes can achieve quite high accuracies of estimates. However, in some situations, accuracy of EBV-s wasn't sufficient to perform accurate selection decisions.

The matrix which indicates these relationships between individuals is called numerator relationship matrix, **A**. Certain assumptions about the relationships have to be made in order to compute this matrix, so for example in the case of full sibs, it is assumed that they have 50% of all alleles identical by descent (IBD) which is not always the case due to Mendelian sampling. Moreover, pedigree is not always complete or it could contain errors, therefore these relationships are even less reliable.

In commonly measured traits, other drawback of this approach is the generally long generation interval which is necessary to collect the progeny data and perform evaluation. In the case of boar taint, information about its indicator compounds, androstenone and skatole could be obtained only in abattoirs on male carcasses, therefore the accuracy is showed to be even lower.

Genomic selection has been introduced for the first time by Haley and Visscher (1998) as a new approach for selecting the superior individuals in animal breeding using information from entire genome. Couple of years later, Meuwissen et al., (2001) suggested a methodology how to perform genomic selection. That methodology required high density marker information which was unavailable at the time due to cost of existing technologies. Subsequently, this became feasible for majority of livestock species with the availability of commercial SNP chips.

In this approach, relationships between individuals are not assumed but calculated from the thousands of available genotypes and used to construct genomic relationship matrix, **G**.

Incorporating the information of thousands of markers simultaneously throughout the whole genome into breeding scheme, accuracy is improved (Van Raden, 2009) since that captured information could be very precise and shared through common ancestors earlier than in the known pedigree (Van Raden, 2009).

Given the specific properties of causing chemical compounds described and practical issues related to boar taint, classical selection approach failed to offer solution. If the proper methodology and newly discovered information of genomic markers is utilized, application of genomic selection could allow positive solution in breeding programs against tainted boars worldwide, which is the main subject of this thesis.

2. LITERATURE REVIEW

2.1. ECONOMIC IMPORTANCE OF BOAR TAINT REDUCTION IN PIG PRODUCTION

Advantages of uncastrated male pig production

In pig production, castration was used for a long time mainly to avoid the undesirable property of tainted carcasses, as described in the introduction. However, if the efficient solution against boar taint became practised in the future, several benefits, mainly economical will support the production of uncastrated males. Advantages associated with the production of boars compared to castrates are mainly related to improved feed conversion with (up to 9 %), higher daily gains (up to 14 %) and higher leanness (up to 20 %) in non castrates (Meat and Livestock Commission, 1989; Babol and Squires, 1995; Bonneau and Squires, 2004). Additionally, if the surgical castration is not performed, possible complications as well as veterinarian costs could be reduced. Therefore, from a commercial standpoint, this system of pig production is economically beneficial. In certain situations, methods used for estimation of lean meat percentage in carcasses could underestimate carcass value of entire males since the calibration is performed on national pig populations, which include both castrates and gilts. Nevertheless, Anderson et al., (1997) found that gross margin per pig place and year was slightly lower for entire males when compared to castrated pigs, with differences being insignificant when based on the Swedish grading system. However, the difference was significant and in favour of entire males when based on dissected lean. In order to avoid this problem, national evaluations should be updated.

Some studies estimate that the production of entire males could increase profit by $5/pig (de Lange and Squires. 1995), or increase the profitability of pig production by 30 % (Lin et al., 2006). These results can greatly improve the world's pork production, which was 112 million tonnes in 2012 (FAO Biannual report on global food markets, 2013) and with further growth expected in the coming period.

In addition to increased production, entire males have improved some meat quality traits. Miyahara et al., (2004) have found higher values of redness and better water holding capacity in meat from entire males than castrates. Consumers may find more acceptable and view as an advantage (EFSA Report, 2004) a significant increase in the proportion of muscle tissue as well as lower fat content, about 5 %, (Babol and Squires, 1995) with more unsaturated fatty acids.

Disadvantages of the entire male pig production

However, production of entire male pigs is also associated with certain disadvantages that should be mentioned. Besides the boar taint as a major issue, problems with this way of production usually occur during the final stage of fattening and mostly relate to aggressive behaviour and fighting of animals (Bonneau and Squires, 2004). When boars reach maturity they become aggressive which often results in the appearance of bruises, skin abrasion, and sometimes DFD (Dark Firm Dry) and PSE meat (PSE - Pale soft exudative) (Andersson et al., 2003). Both DFD and PSE are undesirable characteristics associated with poor meat quality. The incidence of DFD and PSE meat of boars is not common (Babol and Squires, 1995), but additional attention is required while grouping the animals on farms with avoidance of frequent mixing with other groups and manipulation of animals before slaughter. The general recommendation is to separate the animals by sex, especially if the fattening is planned for heavier slaughter weights (Xue and Dial, 1997). Another and specific

disadvantage related to meat quality is a very low amount of fat which could be unsuitable for meat processors as they consider a certain amount of fat content as vital, and therefore describe the aforementioned fat as "too soft" (Bonneau et al., 1979; Wood and Enser, 1982). Also, a possible outcome of low fat content could be substantially harder texture of meat which further exacerbates the present problem of lack of fats resulting from intensive selection (Burkett, 2009). It is a well known fact that the fats are the most important precursors of aroma flavour characteristics, since most chemical compounds are soluble in fats (Ba et al., 2012). As a result, juiciness which is a trait important for consumers could be degraded. Moreover, in the production of entire males dressing percentage is reduced by 2.5 %, because of the genital contribution in total live weight (Sather et al., 1992).

Bearing in mind previously mentioned studies and the EFSA report (EFSA, 2004) it could be concluded that the major problem associated with production of entire males is boar taint, while other problems regarding meat quality are less important.

2.2.CHEMICAL COMPOUNDS AND CAUSES OF BOAR TAINT

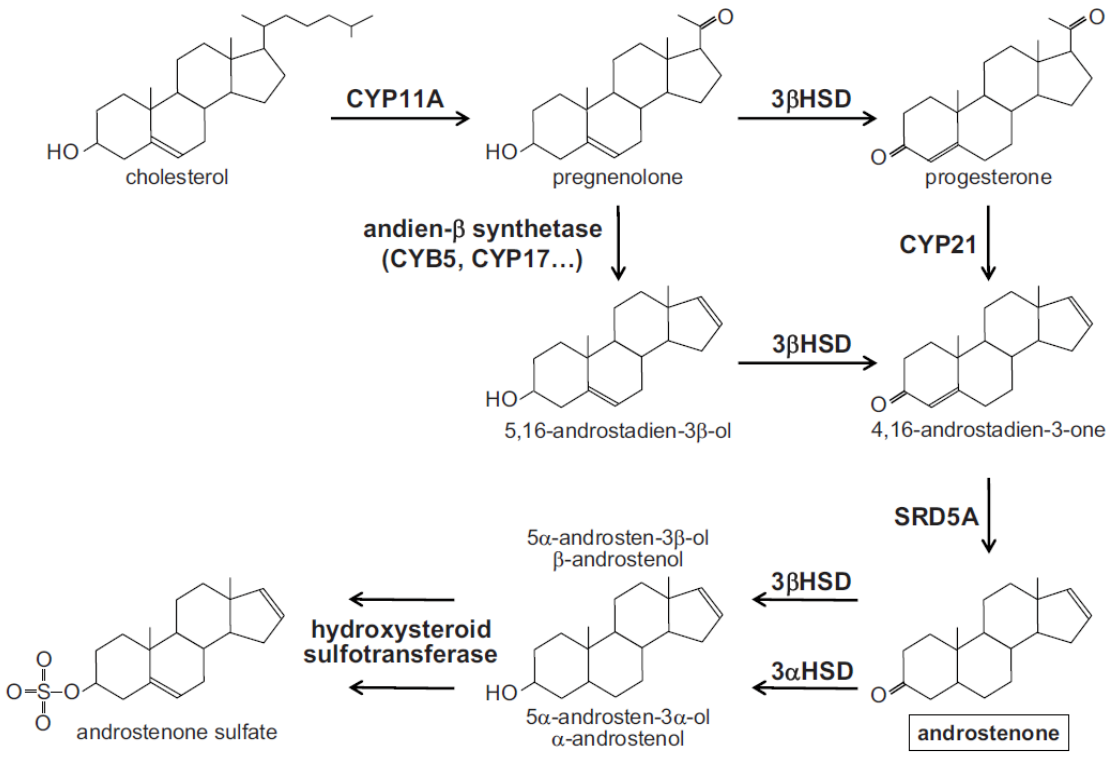
The main chemical compounds responsible for boar taint incidence are androstenone and skatole. Some studies (Moss et al., 1993; Squires and Bonneau, 2004) showed that other compounds could have an effect on boar taint (such as androstenols, indols, etc.), but due to their minor importance determined by numerous studies (Patterson, 1968; Malmfors and Andresen, 1975; Hansson et al., 1980; Bonneau et al., 1992; Xue et al., 1996), this research will focus mainly on androstenone and skatole.

Androstenone

The main chemical compound that is considered responsible for an unpleasant odour and taste of meat from intact boars is androstenone or 5α-androst-16-en-ol. Androstenone is a steroid hormone produced in the male gonads (testes) from commencement of sexual maturity that causes a urine-like odour and taste in meat. Like other steroids, it is synthesized from cholesterol through prognenolone, and then stored in the adipose tissue. Synthesis and degradation of androstenone is shown in Scheme 2 (Robic et al., 2008, cit. Brooks and Pearson 1986). The Scheme 2 shows that the metabolism of androstenone is divided into two stages; hydrogenation and sulfation (Doran et al., 2004; Sinclair and Squires 2005). Excessive accumulation of androstenone in adipose tissue occurs due to the increased synthesis in the testes (Claus et al., 1994) during the phase of early sexual maturity. High levels of androstenone can also be attributed to reduced degradation in the liver or reduced metabolism in testicles. Pigs normally weigh between 100-130 kg when they reach puberty, which is around the age of 5-6 months, at which time it is expected that the level of androstenone in fat tissue will increase significantly (Brennan, 1986). Acceptable concentrations of androstenone in fat are lower than 0.5-1.0 μg/g (Claus et al., 1994). Levels of androstenone and other steroids are also measured in the plasma, but mainly to determine the sexual maturity of pigs and not to predict boar taint as it is known that sexually mature boars may have high levels of androstenone accumulated in adipose tissue and at the same time have normal levels of steroids in plasma (Bonneau et al., 1987).

Studies on consumer sensitivity to androstenone have revealed significant genetic influence (Wysocki and Beauchamp, 1984), and showed that 50% of people are not able to smell androstenone (Xue and Dial, 1997; Blanch et al., 2012). On the one hand, some studies found that 15% of people prefer this kind of smell and taste, while on the other, 35% of people find

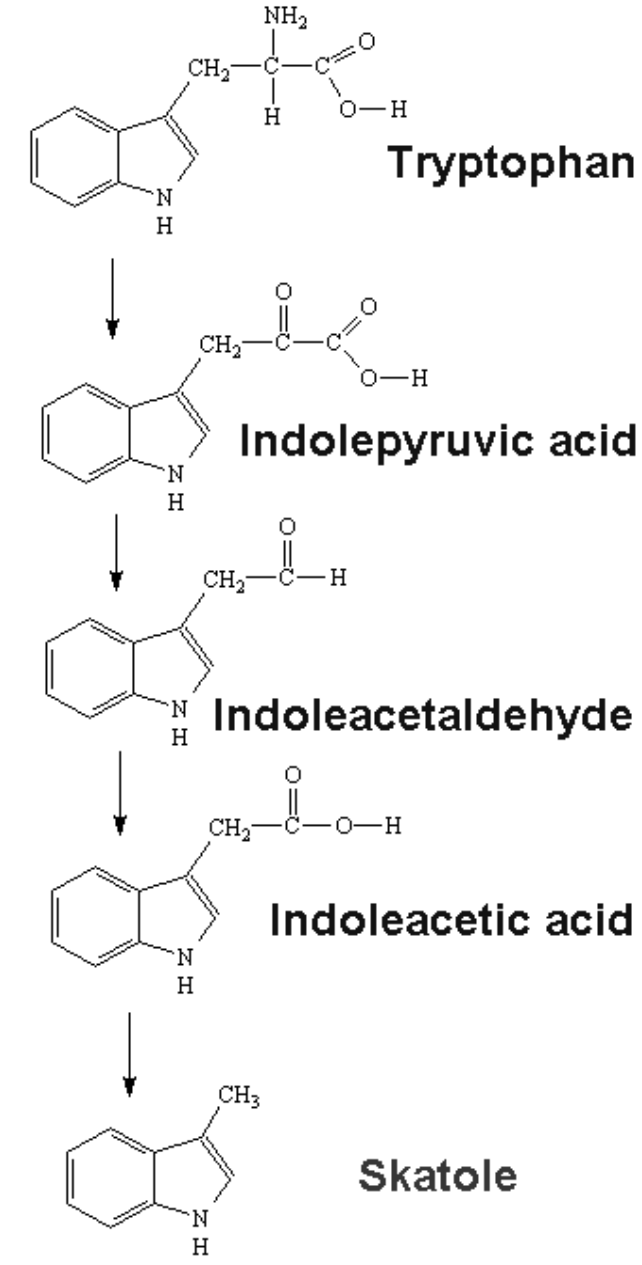
it extremely uncomfortable (Xue and Dial, 1997). A significant impact of gender and origin of the examined people is shown. Women are generally more sensitive to androstenone and only 24% of women cannot detect it, compared to 46% of men (Xue and Dial, 1997). Generally the people from the United Kingdom are less sensitive to boar taint when compared to people from the other European countries (Bonneau et al., 2000). The Americans have approximately the same sensitivity to boar taint as Asian populations, who have shown higher sensitivity compared to the Europeans (Gilbert and Wysocki, 1987).



Scheme 2. Synthesis and degradation of androstenone
(source: Robic et al., 2008, cit. Brooks and Pearson 1986)

Skatole

Skatole or 3-methyl indole is produced by the intestinal bacteria in the colon and causes the smell of faeces or naphthalene (Squires and Bonneau, 2004). It occurs as a derivative of tryptophan primarily affected by the intestine metabolism. In humans and some domestic animals, like goats and cattle, it acts as a pneumotoxin, as opposed to pigs where sensitivity to skatole has not been detected (Robic et al., 2008). The physiological function of skatole still remains unknown in pigs, and given that pigs show no sensitivity, it is assumed that the metabolism of skatole in pigs is specific (Yost, 1989). Metabolism can be divided into two phases; the activation of enzymes from the cytochrome P450 group, and sulfoconjugation (Robic et al., 2008). In the first phase, the key step of skatole degradation is performed by hepatic cytochrome P450. This enzyme belongs to a group of monooxygenase enzymes found in the liver whose main role is the oxidation of organic compounds or degradation of xenobiotics, steroid hormones, lipids and other toxic compounds (Nelson et al., 1996). Xenobiotics (Greek. *Xenos* = side) are foreign substances in the body found in excess in certain parts of the cells or in places where they shouldn't normally appear. The most common are ethanol, acetone, pyridine, etc. While androstenone is characteristic of male pigs only, skatole is produced by female pigs as well and why it accumulates more in fat tissue of boars at the time of sexual maturity than in gilts is still unclear (Squires and Bonneau, 2004). Acceptable levels of skatole concentrations in fat tissue are <0.25 μg/g (Mortensen et al., 1986). Unlike androstenone, where insensitivity or anosmia of people to boar taint exists only within certain groups of people, extensive international research (Bonneau et al., 2000) has confirmed that all people are highly sensitive to skatole. Therefore, it is recommended that priority should be given to lowering skatole levels.



Scheme 3. Skatole synthesis from tryptophan by intestinal bacteria activity (source: http://www.chm.bris.ac.uk/motm/skatole/mechanism.gif)

Methods for androstenone, skatole and boar taint detection

For a long time in the scientific and professional literature, methods for detecting boar taint and its associated components have not been consistent, which made it difficult to interpret and compare results between research studies. Recent reviews of methods for boar taint detection using sensory evaluations and identification of the chemical compounds, androstenone and skatole by laboratory methods or rapid detections are presented by Font-i-Furnols (2012), Haugen et al., (2012) and Aluwe et al., (2012).

Important factors in setting up a sensory evaluation study of boar taint are selection of product samples (slices of fresh meat/fat, sausage, smoked sausage, bacon, salami, etc.); the area where the test will be performed (room/hall, home of consumer, shopping mall, etc.); sensory profile of evaluators (gender, age, origin, etc.) and the method of heating the sample and temperature (heated plates, microwave/oven, pan, etc.; 70 - 250C°). 57 published scientific papers in the last 28 years from the Font-i-Furnols (2012) points out that it is very difficult to make general conclusions regarding the sensory methodologies, and emphasizes further need for harmonization and standardization of threshold levels for androstenone and skatole in order to make better comparisons between studies.

Samples of adipose tissue for laboratory detection of androstenone and skatole are usually taken at abattoirs and stored in freezers on -20° C in the laboratory. It is not necessary to immediately freeze samples at the slaughterhouse because of the high stability of androstenone and skatole. Usually, adipose tissue is sampled in the neck area without precisely defined locations, although some studies have shown that the location of sampling may have an effect (Haugen et al., 2012). Critical phase in every protocol are sample preparation and extraction, as androstenone and skatole are easily soluble in fats. Most of the studies used liquid-liquid extraction. The review by Haugen et al., (2012) describes the

methods and protocols for laboratory measurement of androstenone and skatole. Of the laboratory methods, the most important, and most frequently implemented are radioimmunoassay (RIA) for androstenone and fluorimmunoassay (FIA) for skatole. Methods for the simultaneous determination of androstenone and skatole are enzyme immunoassay (ELISA) and chromatographic methods such as high-performance liquid chromatography (HPLC) and gas chromatography (GC).

Rapid detection methodology such as colorimetric (Mortensen and Sørensen, 1984) was used at the slaughterlines in Denmark. However, the disadvantage of this method is that it only detects skatole. Among other fast methods worth mentioning are gas detection using electronic noses, hot wire, and cooking tests that were not widely applied due to the various technical limitations.

2.3.GENETICS OF ANDROSTENONE AND SKATOLE

The values of heritability (h^2) for androstenone range from medium to high, 0.25-0.81 in the Danish Landrace and 0.61-0.87 in the Large White, (Xue and Dial, 1997) and imply a significant genetic influence on the androstenone levels in adipose tissue. Heritability for skatole according to research of Tajet et al., (2006) ranged from 0.23 in the Duroc to 0.55 in Landrace. The study of Grindflek et al., (2011) showed somewhat higher values of heritability for androstenone detected in fat tissue and positive correlations of androstenone in fat with skatole levels in fat of Duroc breed (Table 2).

Table 2. Residual correlations (above the diagonal), heritabilities (on the diagonal), and genetic correlations (below the diagonal) for investigated traits in Duroc (Grindflek et al., 2011)

TRAIT	AndroP	AndroF	Skat	Indo	Testo	Esulf	Ediol
AndroP	**0.56 ± 0.11**	0.52 ± 0.11	0.28 ± 0.11	0.33 ± 0.10	0.50 ± 0.08	0.71 ± 0.08	0.68 ± 0.08
AndroF	0.91 ± 0.05	**0.67 ± 0.10**	0.52 ± 0.11	0.56 ± 0.10	0.33 ± 0.11	0.41 ± 0.15	0.43 ± 0.15
Skat	0.44 ± 0.15	0.33 ± 0.14	**0.37 ± 0.09**	0.54 ± 0.06	0.11 ± 0.09	0.16 ± 0.13	0.28 ± 0.13
Indo	0.38 ± 0.17	0.32 ± 0.16	0.71 ± 0.12	**0.27 ± 0.08**	0.15 ± 0.08	0.22 ± 0.12	0.27 ± 0.12
Testo	0.90 ± 0.08	0.80 ± 0.10	0.60 ± 0.17	0.46 ± 0.20	**0.32 ± 0.09**	0.40 ± 0.10	0.65 ± 0.07
Esulf	0.89 ± 0.04	0.83 ± 0.06	0.58 ± 0.13	0.53 ± 0.16	0.83 ± 0.09	**0.64 ± 0.11**	0.60 ± 0.11
Ediol	0.92 ± 0.04	0.83 ± 0.06	0.35 ± 0.15	0.34 ± 0.17	0.93 ± 0.05	0.83 ± 0.06	**0.65 ± 0.11**

[1]SE of estimates included.
[2]AndroP = androstenone in plasma; AndroF = androstenone in fat; Skat = skatole; Indo = indole; Testo = testosterone; Esulf = estrone sulfate; Ediol = 17β-estradiol.

Many studies (Grindflek et al., 2011; Duijvesteijn et al., 2010; Gregersen et al., 2012; Robic et al., 2011; Le Mignon et al., 2010; Tajet et al., 2006) have shown high values of heritability within breeds as well as high variability in levels of androstenone and skatole in fat tissue between breeds. It was found that 5-8 % of Hampshire, Yorkshire and Landrace boars had

high values of androstenone in adipose tissue, while for Duroc boars that number was much higher, around 50% (Zamarataskaia et al., 2009).

Skatole levels also vary between breeds. Landrace and Meishan boars often have the highest levels of skatole in fat tissue, while Large White boars and Hampshire have the lowest levels (Pedersen, 1998; Hortos et al., 2000; Doran et al., 2002).

Although androstenone and skatole are physiologically very different, estimates of genetic correlations between them show positive trends with r_g values of 0.36 for Landrace, 0.62 for Duroc (Tajet et al., 2006) and 0.22 for commercial hybrid pigs 0.22 (Engelsma et al., 2007).

Table 3. Residual correlations (above the diagonal), heritabilities (on the diagonal), and genetic correlations (be- low the diagonal) for investigated traits in Landrace (Grindflek et al., 2011)

TRAIT	AndroP	AndroF	Skat	Indo	Testo	Esulf	Ediol
AndroP	**0.47 ± 0.08**	0.42 ± 0.08	0.14 ± 0.09	0.10 ± 0.09	0.59 ± 0.04	0.76 ± 0.04	0.83 ± 0.03
AndroF	0.98 ± 0.03	**0.49 ± 0.08**	0.32 ± 0.08	0.19 ± 0.08	0.20 ± 0.06	0.40 ± 0.08	0.40 ± 0.08
Skat	0.44 ± 0.12	0.32 ± 0.13	**0.41 ± 0.08**	0.62 ± 0.04	0.12 ± 0.06	0.19 ± 0.09	0.26 ± 0.1
Indo	0.57 ± 0.11	0.50 ± 0.12	0.78 ± 0.07	**0.34 ± 0.07**	0.09 ± 0.06	0.10 ± 0.09	0.18 ± 0.09
Testo	0.93 ± 0.16	0.95 ± 0.21	0.42 ± 0.24	0.66 ± 0.24	**0.07 ± 0.05**	0.48 ± 0.05	0.71 ± 0.04
Esulf	0.89 ± 0.04	0.91 ± 0.04	0.33 ± 0.13	0.51 ± 0.12	0.93 ± 0.18	**0.50 ± 0.08**	0.71 ± 0.05
Ediol	0.88 ± 0.04	0.90 ± 0.04	0.11 ± 0.14	0.26 ± 0.14	0.80 ± 0.16	0.85 ± 0.04	**0.52 ± 0.09**

[1]*SE of estimates included.*

[2]*AndroP = androstenone in plasma; AndroF = androstenone in fat; Skat = skatole; Indo = indole; Testo = testosterone; Esulf = estrone sulfate; Ediol = 17β-estradiol.*

Candidate genes and markers for androstenone

Genetic research on androstenone (Moe et al., 2009; Grindflek et al., 2010; Robic et al., 2011) was mainly focused on the search for candidate genes based on the functional approach (which is expected to find genetic markers within genes of known function) that is, finding the responsible genes and QTLs related to the synthesis and degradation of androstenone. Also, special attention was given to the measurement of other steroid components, which could have a significant impact on important reproductive traits.

Genes related with androstenone synthesis

Synthesis of androstenone and other steroids is controlled by the neuroendocrine system (primarily luteinizing hormone LH) which is influenced by gonadotropin - releasing hormone - GnRH (Zamarataskaia et al. 2009). Androstenone synthesis (shown in Scheme 2) begins with hydrogenation of cholesterol with CYP11A enzyme as a catalyst and produces prognenolone (Robic et al., 2008).

CYP11A

In the pig genome CYP11A is located on chromosome SSC7 and belongs to the family of cytochrome P450 enzymes, monooxygenases which play a role in a number of oxidative conversions of steroids, fatty acids and xenobiotics as mentioned in the previous chapter (Nelson et al., 1996). In the research of Quintanilla et al., (2003) no association was found of this potential candidate gene with levels of androstenone in adipose tissue, while some studies have indicated a possible association between the trait and the location of this candidate gene (Greger, 2000). A study on gene expression (Robic et al., 2011) in Large White boars suggests that variations in the gene CYP11A have no effect on the levels of

androstenone, while the research of Moe et al., (2007) found higher expression of this gene in Landrace and Duroc boars with higher levels of androstenone, but this had no statistical significance.

CYP17 and CYB5

The next reaction shown in Scheme 2 is catalysed by two enzymes (Meadus et al., 1993) from the same family as cytochrome P450, c17 (CYP17) located on chromosome SSC14 and cytochrome b5 (CYB5) located on chromosome SSC1. For the CYP17 gene no association has been found with androstenone and skatole levels (Lin et al., 2005). *In vitro* studies of gene expression of CYB5 in various tissues have shown a high association of "G" to "T" substitution at the location 8 bp from the 5' end on SSC1 with the low levels of androstenone in adipose tissue (Lin et al., 2005). Davis and Squires (1999) detected a positive correlation of the CYB5 protein and CYB5 mRNA with levels of androstenone in the fat tissue. However, numerous other studies have not identied mutations as being associated with levels of androstenone in adipose tissue (Quantilla et al., 2003; Lee et al., 2005; Duijvesteijn et al., 2010; Grindflek et al., 2011). Still, it remains a potential candidate gene (Robic et al., 2008).

CYP21

The gene CYP21 also belongs to the family of cytochrome p450 and is considered as a candidate gene due its location, within the SLA (Swine Lymphocyte Antigen) complex between markers LRA1 and S0102 on chromosome SSC7 (Arasta et al., 2007; Payne and Hales, 2004). In an Australian study (Arasta et al., 2007) segregation analysis of 36 SNPs outside the coding region and 14 within the coding region showed no association with levels of androstenone, and consequently the authors dismissed this gene as the QTL on

chromosome SSC7. The extensive genomic association study of Moe et al., (2009) on the Norwegian Landrace and Duroc boars showed no associations for SNPs within introns, but showed an association of three SNPs with skatole concentrations within Duroc. Therefore, additional research needs to be conducted for this candidate gene.

Genes related with androstenone degradation

Degradation of androstenone begins with the reaction where HSD3B hydroxysteroid-dehydrogenase (3α-HSD and 3β-HSD) enzymes catalyse degradation of androstenone to α-androstenol and β-androstenol with the NADH as a cofactor (Doran et al., 2004). Degradation of androstenone occurs in the testes and liver, although the latter is responsible for the majority of the degradation. The subsequent reaction is catalysed with the enzymes hydroxysteroid – sulfotranspherase (SULT2A1 and SULT2B1) and UDP-glucuronosyltransferase (UGT), which mediate the further decomposition to androstenone-sulfate (Zamarataskaia et al., 2009).

HSD3B (3α-HSD and 3β-HSD)

HSD3B gene is located on chromosome SSC4 at a location near the detected QTL and has been considered as a potential candidate gene (Robic et al., 2008). In the study of Moe et al., (2009) a SNP at the location of the NGFI-B (Orphan nuclear receptor) gene was identified which has the function of transcriptional regulation of the 3β-HSD enzyme. This polymorphism had a significant effect on the levels of androstenone in adipose tissue without affecting other steroids, therefore becomes a potential genetic marker for selection for low androstenone. Also, Grindflek et al., (2011) found a further potential candidate gene, UXS1 on chromosome SSC3 which encodes UDP xylose, an enzyme essential for the formation of progesterone with effect on the 3β - HSD enzyme.

SULT2A1 and SULT2B1

Genes responsible for the SULT2A1 and SULT2B1 enzymes in the subsequent reactions of androstenone degradation are not precisely located on the genomic map of pigs. Some enzyme expression studies (Sinclair et al., 2006) showed that the animals with high concentrations of androstenone in fat tissue had significantly lower SULT2A1 enzyme activity in testis and liver. This was not confirmed by other studies, thus the possible association of this enzyme and androstenone concentrations needs further research. Recent study of Moe et al. (2014) has shown different expressions of SULT2B1 enzyme between Norwegian Landrace and Duroc boars in testes and liver, and they suggested that higher concentrations of androstenone in Norwegian Landrace are result of low degradation in first metabolic phase, while in Duroc in second metabolic phase.

Table 4: List of candidate genes involved in synthesis and degradation of androstenone

Candidate gene	Location	Physiological role	Reference
CYP11A	SSC 7	Androstenone synthesis-hydrogenation of cholesterol to produce prognenolone	Greger, (2000); Robic et al., (2011); Moe et al., (2007)
Cytochrome c17 (CYP17)	-	Androstenone synthesis	Lin et al., (2005)
Cytochrome b5 (CYB5)	-	Androstenone synthesis	Davis and Squires (1999); Grindflek et al., (2011); Duijvesteijn et al., (2010); Quintanilla et al., (2003); Lee et al., (2005)
CYP21	SSC7	Androstenone synthesis	Arasta et al., (2007); Moe et al., (2009)
HSD3B hydroxysteroid-dehydrogenase	SSC4	Androstenone synthesis	Moe et al., (2009)
(SULT2A1 and SULT2B1)	-		-
UDP-glucuronosyltransferase (UGT)	SSC3	Androstenone synthesis	Grindflek et al., (2011)

Candidate genes and markers for skatole

Synthesis of skatole begins in the colon by intestinal bacteria activity which degrades amino acid tryptophan into skatole. Complete degradation of tryptophan is limited by anaerobic conditions in the digestive system, especially in the initial phase, ultimately resulting in excessive accumulation of skatole as the end product (Wesoły et al., 2012). Of the total skatole production in the intestine, approximately 87 % will be absorbed through the intestinal walls and transferred to the liver by blood (Xue and Dial, 1997). The remaining 13% of produced skatole is excreted from the body through faeces. Differences in the concentrations of skatole in faeces were not noted between sexes, while higher levels of skatole in fat tissue occur only with some male pigs. Studies have shown that various groups of bacteria, most commonly of the genus *Clostridium* and *Escherichia coli*, can degrade tryptophan to indoleacetate by the first three steps (shown in Scheme 3) whereas the final step is mainly catalysed by bacteria of the genera *Clostridium, Lactobacillus* and *Bacteroides* (Zamarataskaia et al., 2009, Wesoły et al., 2012). Furthermore, in the first three steps, enzymes from the cytochrome P450 group perform a major role in addition of the hydroxyl group, while in the second and final step the main role of sulfoconjugation reaction is performed by sulfotransferase enzyme.

Table 5: List of candidate genes involved in degradation of skatole

Candidate gene	Location	Physiological role	Reference
CYP2E1	SSC 14	Skatole degradation	Rowe et al., (2014); Grindflek et al., (2011); Moe et al., (2009); Skinner et al., (2005)
CYP21	-	Skatole degradation	Moe et al., (2009)
CYP2A6	-	Skatole degradation	Lin et al., (2004)
SULT1A	SSC3	Skatole degradation	Lin et al., (2003)

CYP2E1, the most important candidate gene in skatole metabolism

Enzyme isoforms from the family of cytochrome P450, CYP2E1 and CYP2A catalyse most reactions and the largest number of studies have emphasized the primary role of CYP2E1 located on chromosome SSC14 as the most important in the metabolism of skatole (Babol et al., 1998; Diaz and Squires, 2000; Le Mignon et al., 2010, Terner et al., 2006; Tajet et al., 2006; Rowe et al., 2014). The high activities of these enzymes are commonly associated with low levels of skatole in adipose tissue (Squires and Lundstrom 1997; Babol et al., 1998; Doran et al., 2002; Lin et al., 2006; Zamarataskaia et al., 2009). Skinner et al., (2005) found six SNPs in the coding region of this gene, where one (AJ697882_2412) was associated with higher levels of skatole. The latter study was performed on the Danish commercial hybrids, so the possibility remains that it was the influence of the breed. A study by Zamarataskaia et al., (2006) has shown that the both male and female pigs had the same activity of CYP2E1

until they reached the weight of 90 kg, while in male pigs activity was decreased at a weight of 115 kg. A potential insight into the complex interaction of CYP2E1 and androstenone is given by studies of Tambyrajah et al., (2004) and Doran et al., (2002) who found that androstenone reduces the CYP2E1 promoter activity by inhibiting the binding of transcription factors (HNF-1 and COUP-TF1). Furthermore, the research by Moe et al., (2009) and Grindflek et al., (2011), found an association between SNPs and haplotypes within the region of CYP2E1 with skatole levels without affecting the levels of androstenone in Duroc and Norwegian Landrace, which strengthens further this gene as a potential candidate.

CYP2A

Another important enzyme in the physiology of skatole is CYP2A (Banoglu et al., 2001; Diaz and Squires 2000), which also participates in the catalytic reactions of degradation. According to Zamarataskaia et al., (2009), the activity of the enzyme CYP2A is similar to CYP2E1 with some differences determined in their expressions mostly related to nutrition. The location of the gene for this enzyme, just as some other enzymes from the P450 family, is not pinpointed precisely on the pig genome. This has prevented their significant and consistent effects on the levels of skatole in fat tissue to be confirmed.

SULT1A

The second and final step in skatole degradation is the reaction of sulfoconjugation performed by enzyme sulfotransferase SULT1A. Although the study of Lin et al., (2004) identified a mutation on 546 base (A => G) within the coding region of SULT1A1, which significantly reduces its catalytic activity, the research by Skinner et al., (2006) showed no association of this polymorphism with levels of skatole in fat tissue. Therefore it is not considered significant in the case of boar taint.

2.4. POSSIBILITIES OF GENETIC SELECTION METHODOLOGIES

Classical selection approach in animal breeding and breeding against tainted boars

The classical selection approach utilises pedigree and phenotypic information to predict breeding values of individuals. In this approach, pedigree information is included typically with the assumption of an additive infinitesimal model through the numerator relationship matrix **A** (Fisher, 1918; Wright, 1921) which is equal twice the matrix of kinship coefficients between individuals. In animal breeding, this approach showed huge success for the majority of economically important traits. For example, the annual milk production of Holstein dairy cow has increased by 110 kg per animal while in pig production, the feed conversion ratio has decreased by 50% in the last 50 years (Dekkers, 2004; Eggen, 2012). However, this approach is unsatisfactory in cases where the traits have low heritability or few recordings, if the measurement of traits is only possible late in life (age-limited), on one sex (sex-limited) or in the case of carcass traits or disease resistance traits (Haley and Visscher, 1997; Meuwissen, 2006). The information on Mendelian sampling is also not available at the time of selection if the progeny records are not collected. Moreover, the additional information from candidate genes studies is difficult to include into breeding programs as detected markers do not usually explain sufficient amount of genetic variation.

Boar taint is an example where the classical selection approach has failed to deliver its potential. It is sex-limited, age-limited and difficult to measure as the information is only available after slaughtering the animals. In practice the application of genetic selection has been only partially successful. Due to common physiological pathways and high correlations

of androstenone with other steroid hormones, selection has resulted in low levels of androstenone, but with reproductive problems as a correlated response. Early attempts of genetic selection against boar taint began in the early 80s of the last century (Willeke et al., 1980) when boars ended up with reduced testes size after three generations of selection. Furthermore, the research of Willeke et al., (1987) observed delayed sexual maturity in gilts, and similar results were found by Sellier et al., (1988). Beside the effect on reproductive performance, selection towards high or low androstenone levels after five generations resulted in high unfavourable correlations between growth and androstenone fat concentrations (Willeke and Pirchner, 1989). Sellier et al., (2000) went a step further using a restrictive selection index with two components in the model, the level of androstenone and the size of bulbourethral glands as control of reproductive traits. The idea was to decrease the level of androstenone while preserving optimum male reproduction performance with the bulbourethral glands size. The results showed significantly increased bulbourethral glands but androstenone levels were not reduced, which was explained by the low accuracy of estimated genetic parameters. It is evident from the studies shown that the application of genetic selection in the case of boar taint requires an approach that includes identification of the responsible genetic markers or genomic regions, with special attention given to the accurate estimation of genetic parameters and control of reproductive traits.

Strategies for implementation of genomic information in animal breeding

In agricultural production, selection with DNA markers or MAS (Marker Assisted Selection) was introduced for the first time in the 1990s. Strategies for implementation of marker information were created according to the type of genetic loci: causal mutations or direct

markers and non-functional markers linked to the quantitative trait loci or indirect markers (Andersson, 2001; Dekkers, 2002).

Causal mutations or close markers

In the mid 80s, major genes important for commercial pig production was discovered by Monin and Sellier (1985) which were responsible for low ultimate pH values and poor water holding capacity in some Hampshire lines. The exact location of this dominant mutation in PRKAG3 (RN) gene was later discovered in the study of Milan et al., (2000) what further helped its implementation in breeding programs. The first direct marker implemented in pig breeding programs through MAS was the halothane gene (Fuji et al., 2001) or RYR1, responsible for higher susceptibility to stress induced malignant hyperthermia and associated with high lean meat percentage. This identified gene regulates the vital physiological function of Ca^{2+} ions metabolism in skeletal muscles so it was relatively easy to detect and localise the effect, and then to include that information in breeding. These two examples of MAS using direct markers with deleterious effects were quite efficient mainly because of their strong effect on economically important traits. However, direct causative markers are generally hard to detect (Dekkers and Hospital, 2002) so there are only few examples like these that have been applied in practice.

Non-functional or indirect markers

Indirect markers, opposite to previous ones, are abundantly distributed along the genome while their linkage phase with the quantitative trait loci has to be established in order to use it in selection. In order to utilize information of those markers in selection, two strategies

were used: search for candidate genes (as described for androstenone and skatole) and whole genome search for associated chromosomal regions (Dekkers and Hospital, 2002).

The strategy of searching candidate genes assumes that the mutation in genes involved in the known trait physiology could be responsible for trait variation. The review by Andersson and Georges (2004) showed examples of this approach in animal breeding. However, there are two practical issues concerning this approach (Hayes, 2012), first is the large number of candidate genes that have to be tested and second, causative variants may lie in unknown network components.

The boar taint related compounds are good example. As described in the previous chapter, most of the candidate genes for androstenone and skatole synthesis and degradation didn't explain enough genetic variance and most were not subsequently validated to make reliable decisions in selection.

Another strategy, searching for the polymorphisms on a whole genome scale, utilizes linkage disequilibrium to find associated chromosome regions with a phenotypic trait. DNA polymorphisms or markers, are some segments of DNA that have no coding function, but their allele variation on molecular level could be linked to phenotypic variation of a quantitative trait. Additionally, if those markers are physically close enough to the causative mutation, they could be used in selection. If they aren't close, their association will be present only within families, perhaps with different phases and will be broken down by recombination in the following generations. This approach has been used only rarely in animal breeding using markers such as RFLP (restricted fragment length polymorphisms) but more often using microsatellites. Besides the practical necessity for close linkage between the marker and QTL, a disadvantage in this approach was the low resolution of mapping, so the region that is mapped usually contains too many possible candidate genes

(Andersson and Georges, 2004) and requires detailed fine mapping for reliable implementation. Moreover, a large number of progeny from each family is necessary to test due to the large confidence intervals (Hayes, 2012).

QTL detection studies

There were three notable QTL detection studies on androstenone and skatole using microsatellite markers all reviewed by Robic et al., (2008) and Zamarataskaia et al., (2009). The first study in 2003 was carried out by the French group of authors, Quintanilla et al., (2003) and was focused mainly on androstenone fat concentrations. In their analysis they used three-generation experimental crosses between Large White and Meishan pig breeds genotyped for 137 microsatellites along all autosomes and X chromosome. With two statistical approaches used, they detected genome-wide significant QTLs on chromosomes 3, 7, and 14 that explained 7 to 11%, 11 to 15%, and 6 to 8% of phenotypic variance, respectively. Their results have clearly shown that several genomic regions significantly affected the fat concentrations of androstenone. However, probably the main drawback of the study is the slaughter weight of 80kg, which is considered too low to establish that the boars had reached sexual maturity.

The study of Lee et al., (2004) was carried out by the British group of authors on the same population, crosses between Large White and Meishan pigs. QTLs for fat androstenone were detected on chromosomes 2, 4, 6, 7 and 9. The QTL on chromosome 6 was for fat androstenone and boar taint assessed by the sensory panel. It was expected that the QTLs would be more consistent with previous study, however, only one was consistent on chromosome 7 for androstenone. They also detected a QTL on chromosome 14 for skatole.

The third study analysed QTLs on a Landrace population (Varona et al., 2005) on 10 genomic regions and for each region two or three microsatellite markers were chosen. The

regions were selected based on the published QTLs for growth and fatness, and were similar with the mapped regions from the previous two studies (on chromosome 2, 3, 4, 6, 7, and 9). Based on the results of their study, no QTL for fat androstenone was segregating in investigated population. One QTL was detected for skatole concentrations on chromosome 6, which is in concordance with the study of Lee et al., (2004).

In all three studies described, a lack of consistency is present, as only one QTL was confirmed on a same location for androstenone with the first two studies, and one for skatole with the second and the third study. The reasons could be the small population sizes or even low marker coverage in the third study. Moreover, final weights in the first and the second study were 80 and 85 kg, which is generally considered to be a low mass for proper indication of sexually matured boars.

In order to further enlighten the responsible QTL-s for boar taint compounds, more focus should be given to these mentioned factors.

Besides the above-mentioned studies with microsatellites, several QTL detection studies on androstenone and skatole were made using genome-wide SNPs (Moe et al., 2009; Duijvesteijn et al., 2010; Grindflek et al., 2011a; Grindflek et al., 2011b; Gregersen et al., 2012; Rowe et al., 2014).

Moe et al., (2009) performed a genome-wide search with 275 selected SNP-s on 1102 Duroc and 1726 Norwegian Landrace boars. They were focused on candidate genes for androstenone, steroid hormones, skatole and bulbourethral gland size based on previous QTL literature findings as well as the published microarray results. Their results suggested the that polymorphisms for skatole (CYP2E1, CYP21, CYP2D6, CYP2C49, NGFIB and CTNND1) could be used to reduce levels of boar taint in both breeds without affecting levels of other hormones or bulbourethral gland size.

Study of Duijvesteijn et al., (2010) was focused on androstenone concentrations in fat in Duroc boars. They genotyped 987 boars with 60k Illumina beadchip and revealed major genetic factors on SSC1 and SSC6 showing moderate to large effects. However, it was shown that the larger region on chromosome 6 associated with androstenone includes several candidate genes potentially involved in physiological pathways of other androgens.

The study of Grindflek et al., (2011a) tried to explore the relationship between genetic factors of androstenone and skatole and fertility related traits using the low density 6k Illumina beadchip. They used the same animal material as Moe et al., (2009) and detected 27 regions significant at a genome-wide level ($P < 0.05$) of the which most important were associations in 6 regions affecting skatole and indole on chromosomes 1, 2, 3, 7, 13, and 14; 5 regions on chromosomes 3, 4, 13, and 15 affected androstenone, testosterone, and estrogens and 1 region on chromosome 6 affecting androstenone in plasma without apparent negative effects on testosterone and estrogens. This study however, confirmed that all significant QTLs for fat androstenone also affected sex hormones important for proper fertility-related functions.

The GWAS conducted by Grindflek et al., (2011) on 1,251 purebred Norwegian Landrace and 918 purebred Duroc boars genotyped for 60k Illumina beadchip aimed to detect new QTLs associated with boar taint compounds and related sex steroids. They detected 14 genome wide significant regions for androstenone in both breeds, while 7 of those (SSC 1, 2, 3, 7 and 15) were common in both breeds. Moreover, all 14 regions affected the estrogens in Landrace as well, while in Duroc only 3 does not affect other sex steroids.

Additionally, for skatole, 10 QTLs for Landrace and 4 for Duroc achieved genome wide significance while 4 of these were detected in both breeds. They concluded that since the QTLs for skatole do not negatively affects other steroids it should be preferred over androstenone in breeding schemes.

Gregersen et al., (2012) in their study searched for QTLs from a data of 923 animals comprised from 3 breeds, Duroc, Danish Landrace and Yorkshire and genotyped on the Illumina 60k BeadChip. They have identified 46 chromosomal regions that affect boar taint compounds and the SNPs that were highly associated with them were used to identify haplotypes. Although the sample size was generally small, their results confirmed the majority of previous QTLs whilst adding new candidate genes. Additionally, they observed little or no overlap of QTLs between breeds which further emphasizes the breed differences in the context of boar taint compounds.

The most recent study of QTL detection was the GWAS of Rowe et al., (2014) on the Danish Landrace population. The power to detect a QTL was additionally increased by divergent selection of littermates for skatole concentration; 500 boars with high skatole (>0.3 μg/g fat) were matched with low skatole litter mates (<0.3 μg/g). The results revealed the most significant QTLs were on chromosome 14 for skatole and chromosome 5 for androstenone. The SNP detected for skatole lies within CYP21E gene, which encodes a protein responsible for skatole degradation.

From the review of these notable QTL detection studies, important points should be drawn out. Androstenone and skatole are completely different by their physiological functions and their complex relationship in terms of accumulation in fat tissue at the time of sexual maturity remains unclear. However, their genetic relationship became more or less clearer. Positive genetic correlations and high heritabilites have been confirmed by numerous studies, but still, additional research has to be conducted for overall strategy against boar taint. Detected QTL regions associated with both androstenone and skatole have been rather inconsistent between studies and between breeds. These inconsistencies are probably related to different study designs, fixed effects used in statistical models (age, slaughter weight etc.)

or availability of sufficient data. Therefore, special attention in future research should be given to these factors.

Some of the QTLs primarily for skatole, have been confirmed between these studies without undesirable correlations with reproduction traits, therefore it might be considered in breeding programs (Moe et al., 2009; Rowe et al., 2014).

2.5.GENOMIC SELECTION APPROACH

Considering that the traits of interest in livestock breeding are usually of quantitative character and therefore under the control of a large number of genes that have a relatively small effect on the trait, approach that utilizes all genetic markers throughout the genome called genomic selection has been proposed (Meuwissen et al., 2001; Dekkers, 2004).

The problems present in QTL mapping approach previously described based on linkage could be resolved with a very dense marker maps, because the markers would be physically close to the QTL and probably in population wide linkage disequilibrium (Meuwissen et al., 2001).

The advent of high density single nucleotide polymorphisms (SNPs) beadchips and recent improvement of methods for sequencing the genome (Next Generation Sequencing) has provided the possibility of genotyping animals for thousands of markers so that genomic selection has become a commercially feasible and effective option in agricultural production.

However, the first disadvantage in this approach that arises when genomic information from very dense marker map is used is of statistical nature. Linear regression on markers as a simple method in association studies is not possible in this situation because the number of

SNP effects that have to be estimated was much larger than the number of phenotypic observations collected. Therefore, this approach was infeasible.

Other solutions of genomic selection methodology were proposed and commonly grouped into a) genomic best linear unbiased prediction (GBLUP), and b) Bayesian approaches which assume various priors in which some subsets of markers are assumed to explain more variance than others. The latter are advantageous when the number of QTL explaining the variance is small (Daetwyler et al., 2010) and a number of Bayesian methods have been proposed differing in their assumptions for partitioning SNP into those with 'large' or 'small' effects and the distributional assumptions within these classes.

Genomic BLUP (GBLUP) methodology

Before the advent of dense genomic information, pedigree information was used to calculate the additive genetic relationships for all loci (Falconer and Mackay, 1997) and together with phenotypic performance was used to select superior animals. This was achieved by the estimation of breeding values (EBV) for every individual using BLUP methodology (Best Linear Unbiased Prediction). The method was developed by Henderson (1949) in which breeding values together with fixed effects could be estimated simultaneously (Mrode, 2005). However, the genetic architecture of the trait, the gene location and its effect still remain a subject of intensive research as they are treated as being contained within a "black box" (Dekkers, 2002; Hayes, 2009; Daetwyler et al., 2010; Kemper et al., 2012).

In the genomic BLUP approach (Genomic Best Linear Unbiased Prediction), the pedigree relationship matrix is replaced with the genomic relationship matrix (GRM) which describes genomic relationships between individuals and is calculated from the SNP (Single

Nucleotide Polymorphisms) genotypes (Habier et al., 2013). In that way, all SNPs are treated equally important, e.g. genetic variance is equally assigned to all SNP-s (Clark and van der Werf, 2013; Van Raden et al., 2009). This methodology showed improvement of accuracy compared to traditional BLUP (Daetwyler et al., 2008; Goddard, 2009) since the real genotypes could represent more reliable additive relationships between individuals than expected from pedigree (Meuwissen et al., 2001; Villaneuva et al., 2005). For example, the relationship in the GRM of two full siblings may vary from 0.4 – 0.6 instead of the expected value of 0.5 contained within the numerator relationship matrix A (Lee et al., 2010).

The typical model for GBLUP is defined by the following equation:

$$y = 1_n\mu + \mathbf{Z}g + e$$

where y is a vector of phenotypes and μ is the mean, **Z** is a design matrix allocating records to genetic values, **g** is a vector of additive genetic effects of individuals assumed to be distributed MVN(0, σ^2 **G**), and **e** is a vector of residuals σ^2_e assumed to be distributed MVN(0, σ^2 **I**).

Research of Van Raden et al., (2009) showed that the accuracy of breeding values with genomic predictions was significantly higher than achieved by traditional parent averages. Other researchers also showed improvements of genomic evaluations using GBLUP compared to traditional BLUP (Moser et al., 2009; Harris et al., 2008; Hayes et al., 2009). Moreover, if many QTLs exists each of them with small effect on the trait, this method showed similar accuracies of GEBV-s to Bayesian methodologies, but the latter showed better performance in situations where a limited number of QTL-s strongly effect the trait (Daetwyler et al., 2010).

The study on simulated data by Muir (2007) compared the accuracies achieved by predictions of estimated breeding values based on genome wide markers (GEBV) and classical Best Linear Unbiased Prediction (BLUP) while taking into account other possibilities such as low heritability, number of generations of training, marker density, initial distributions, and effective population size (Ne). Results showed if more training individuals with both genotypes and phenotypes were collected, the accuracy of GEBV was higher than EBV. A further conclusion was the higher accuracies of GEBV-s compared to traditional selection and therefore advantage for traits of low heritability.

Estimation of dominance effects with genomic information

One way of incorporating genomic information in genomic prediction models is through the GRM, as described in the example of GBLUP. The GRM defines the additive genetic covariance between individuals. In genomic relationships, coefficients are estimated more accurately compared to pedigree relationship matrix because the GRM can capture additional genetic variation due to Mendelian sampling (Meuwissen et al., 2001; Villaneuva et al., 2005). However, most of the models are used to estimate additive genetic effects with no specific attention given to the estimation of non-additive effects, such as dominance effects. The reason for this was necessity as a very large number of reference families is required to achieve higher accuracies of estimations and overcome computational difficulties (Misztal et al., 1998). However, with introduction of dense genomic markers, sufficient information can be provided to estimate relationships more precise and specific to genomic regions. When computational barriers are no longer an issue, inclusion of dominance genetic effects in national evaluations becomes worthwhile to consider. For example, the study of

Su et al., (2012) on Duroc pigs showed the additional genetic variation of 5.6% due to dominance genetic effects estimated by GBLUP. Nishio et al., (2014) used PIC pig data that was made publically available and showed improvement of accuracy of GBLUP model when dominance effects were included (expressed as correlation between the estimates and the true values), but the authors claimed that the results could be better with crossbreds as the degree of dominance genetic variance was small in their dataset. It is expected however, much higher dominance genetic variation in crossbred populations compared to purebred populations, therefore, the information from crossbred populations could increase accuracy of genetic evaluation for purebreds (Su et al., 2012).

Bayesian approach in genomic selection

As previously mentioned, in the association studies with a very dense marker maps, number of markers is usually larger than the number of collected phenotypic observations. Because of that, the analysis of multiple regression cannot be used to simultaneously estimate effect for each marker (Meuwissen et al., 2001). However, in Bayesian approach, this problem is solved by setting up prior distributions. The prior is concerned with the distribution of the QTL effects, for example what fractions of QTLs have small effect or strong effect on the trait (Hayes, 2009). Inferences about the variance parameters are estimated from posterior distributions of the estimates which represents the experimenter's belief in the value of the parameter after considering the experimental data and given the prior belief. In addition, it is plausible to create Markov chains in order to construct approximations to the posterior distributions (Garrick et al., 2014). This methodology was introduced by Meuwissen et al., (2001) together with a Markov Chain Monte Carlo (MCMC) algorithm using Gibbs sampler as a tool to sample the marker effects for each locus using a large number of iterations. So

far, based on different prior assumptions, various methods are proposed and tested under the Bayesian framework.

This approach showed good performance and high accuracy of predictions in situations where significant QTLs exist, while on the other hand the same performance as GBLUP in cases where the traits are affected with many SNPs with small effect (Clark et al., 2009; Daetwyler et al., 2009). Many studies showed the superiority of BayesB method, which is the model with mixture distribution where one of the distributions is associated with zero locus effects and other with non-zero effect. Research of Meuwissen et al., (2001) explained that most genomic regions have no QTL-s while only few have certain effect, therefore, this methodology could be very useful if the trait architecture follows assumptions of methodology. However, one of the important drawbacks of this method is the high computational requirements and the time necessary to perform the predictions compared to GBLUP, which is fast and easy method to perform.

Accuracy of genomic prediction

In animal breeding, genetic gain per year is defined by the following formula (Falconer and Mackay, 1996):

$$\Delta G = \frac{(A\sigma_g\, i)}{I}$$

where A is the accuracy of selection; σ_g is the standard deviation of the additive genetic variation; i is the selection intensity and I is the generation interval. Some factors from this equation could be improved, but with limited resources it is difficult to maximize them

simultaneously. For example, σ_g hardly varies, especially within breeds. Accuracy of selection achieved by the classical selection approach in progeny testing schemes is already high and so is the genetic gain, because of the large pedigree populations and records available for selection candidates. However, in cases of low heritable traits, disease resistance traits or traits that can be measured only after slaughtering animals such as boar taint, the desired accuracy of selection cannot be obtained (Meuwissen, 2003). Moreover, the generation interval is also high what further decreases overall gain as the time necessary for selection candidates to reach maturity will increase. Selection intensity could also be increased by lowering the number of animals selected, but on the other side, more attention have to be given to negative effects that could arise (Daetwyler et al., 2007).

The accuracy of calculated GEBVs of individuals is defined as the correlation between the GEBVs and the true breeding values (TBVs) (Falconer and Mackay, 1996).

The availability of phenotypic observations in the training set (large set of genotyped individuals with phenotypes) as well as the heritability of the traits has a considerable effect on the accuracy of GEBVs (Meuwissen et al., 2001; Daetwyler et al., 2008). In order to accurately estimate marker effects for the lowly heritable traits, more phenotypic records need to be collected than for highly heritable traits. Another factor that could have an effect on the accuracy is the number of markers. In the GBLUP approach, all markers are treated equally but with regression based methodologies like Bayes B, number of markers could affect the accuracy.

Therefore, Daetwyler et al., (2010) included that assumption for GBLUP methodology in the following equation:

$$r = \sqrt{N_p h^2/(N_p h^2 + M_e)}$$

where N_p is the number of individuals in the training population; h^2 is the heritability and M_e is number of independent chromosome segments.

For the BayesB method formula was as follows:

$$r = \sqrt{N_p h^2/(N_p h^2 + \min(M_e, N_{qtl})}$$

All of the factors included in these formulas as length of the genome, effective population size and trait architecture represent the main factors that contribute to the accuracy of the GEBV-s. However, other factors may also contribute to the accuracy of prediction like size and the structure of the reference population, marker density and the selected approach used for prediction (Nirea et al., 2011). Nevertheless, all of those factors depending on the situation have to be evaluated before implementation of the acquired results in practice.

Application of genomic information in selection against androstenone and skatole

Boar taint has very specific limitations. It is measurable only after slaughter, observed only in male animals, and there are likely to be biological constraints as androstenone shows positive genetic correlations with reproductive traits. Despite these challenges, with the existing knowledge and the discovery of new genetic markers and QTL's, genomic selection is a viable long term solution in the prevention of boar taint and could become a practical solution in the near future.

Until now, successful application of genomic selection in breeding programs of cattle and sheep is practiced in the U.S., Canada, UK, Netherlands, Australia and New Zealand (Van Raden 2009; Spelman et al., 2010; Duchemin et al., 2012; Ibañez-Escriche and Gonzalez - Recio) with recent applications in pigs and poultry (Wellman et al., 2013; Wang et al., 2013). Genetic gain is largely dependent on the amount of genetic variation relevant within a given population or cohort. The studies described previously, indicate that both androstenone and skatole, have moderate to high heritabilities and are positively correlated. Combined with the fact that these are sex limited and difficult to measure traits, this provides a strong basis for potential implementation of genetic selection. Unsuccessful or partially successful examples of genetic selection against boar taint described in introduction used classical or conventional approaches that were based on the estimation of genetic parameters with pedigree information. It is known that this approach can be sub-optimal in the case of traits measurable only after slaughter, later in life or if the traits are related to only one gender (Meuwissen 2003; Ibanez-Escriche and Gonzalez-Recio 2011).

Evidence for genetic variation in androstenone and skatole concentrations in fat tissue has been reported in numerous studies between breeds (Grindflek et al., 2011; Duijvesteijn et al., 2010; Gregersen et al., 2012; Robic et al., 2011; Le Mignon et al., 2010). Within breed estimates of heritability range from 0.25 to 0.88 for androstenone, and 0.19 to 0.54 for skatole, reviewed by Robic et al., (2008). However, exploiting this variation is challenging as the trait is age-limited, sex-limited and destructive: only males express taint, it is not expressed until after sexual maturity, and can only be measured after slaughter (excluding invasive techniques). One approach to overcome all these challenges is the use of genomic predictors, available from birth in both sexes and, with adequate training data, capable of delivering high accuracy. Such predictors may either be based upon a handful of causative

mutations explaining a high proportion of the variance, or - via genomic evaluation (Meuwissen et al., 2001).

There has been little consensus in the literature regarding the genetic architecture of boar taint. QTL mapping studies and GWAS appear to identify QTLs that differ markedly by location and effect (Quintanilla et al., 2003; Lee et al., 2005; Grindflek et al.,2011; Rowe et al., 2014). The reason for this may be the different breeds that were used, or this could indicate that many genes have an effect. For androstenone and skatole, major candidate gene that explains larger proportion of genetic variance is not found although some potential candidates are detected for skatole in some breeds what addresses additional research (Rowe et al.2014). The genetic architecture influences the effectiveness and accuracy of different methods of genomic evaluation (Daetwyler et al., 2010) therefore, in the case of boar taint which is influenced by two chemical compounds with different physiological pathways, genomic evaluation methodology should be properly selected.

It has been shown that the androstenone concentration in fat is highly correlated with reproductive traits as well being dependent on the social dominance or litter (Giersing et al., 2000), what suggests non-additive genetic effects on androstenone concentrations. Furthermore, the utilisation of dominance effects could be more efficient in the crossbreeding programs and highly useful for commercial production where the large dominance genetic variation is expected.

3. AIM OF THIS RESEARCH

The main aim of this thesis is to overcome mentioned drawbacks of classical selection theory by using genomic selection methodology on boar taint related compounds, androstenone and skatole, and to provide additional knowledge necessary for long-term solution against boar taint in pig production. Therefore, the targets are:

1. to evaluate GBLUP and five Bayesian methods by testing their accuracies of prediction breeding values for androstenone and skatole
2. to determine which of the proposed methodologies provide most accurate predictions regarding the trait architecture
3. to recommend the best solution in application of genomic information against androstenone and skatole concentrations in Danish Landrace population
4. to evaluate dominance genetic effects using genomic information
5. to provide improved accuracy and unbiasedness of genomic predictions

HYPOTHESIS

Methods of genomic selection, GBLUP and regression based methodologies provide accurate methods for prediction of genomic breeding values for skatole and androstenone.

4. MATERIAL AND METHODS

Animals

All the animals involved in this study were raised under conventional pig production conditions and were not subjected to any experimental procedures. All the samples for the study were collected post-mortem in a commercial abattoir.

Sample collection

Samples were collected at the abattoir from 6,178 entire male Danish Landrace pigs of known pedigree and known farm of origin.

Two samples of adipose tissue were collected from each animal at the abattoir; the first one immediately after the carcass was cut into two sides and the second one an hour later. The former samples were assayed immediately for skatole levels in-house at the abattoir. The second adipose samples and a muscle sample from each animal were stored at -20°C for future analyses.

Selection of animals for genotyping and androstenone analysis

Skatole concentrations (μg/g fat tissue) were analysed using a spectrophotometric method (Møller and Andersen, 1994) and used to select 464 animals with high skatole concentrations in fat tissue (≥0.3 μg/g). These identified animals were then matched to a litter mate with low skatole concentration where available, which was possible in 421 of the cases, and an additional 56 animals with low skatole concentrations were also selected. The priority in the selection process was to pick a pair within the same litter with the highest and the lowest concentration. Therefore, of the 941 animals included in analyses, there were 842 animals

from sib pairs, each pair having an animal with a high and a low skatole concentration, 40 unrelated animals with high skatole concentrations, and 56 unrelated with low skatole concentrations. Distribution of skatole concentrations are shown in *Figure 1.* and *Figure 2.* These 941 entire males had been bred from 128 sires and 441 dams in 441 litters and had been reared on 14 farms.

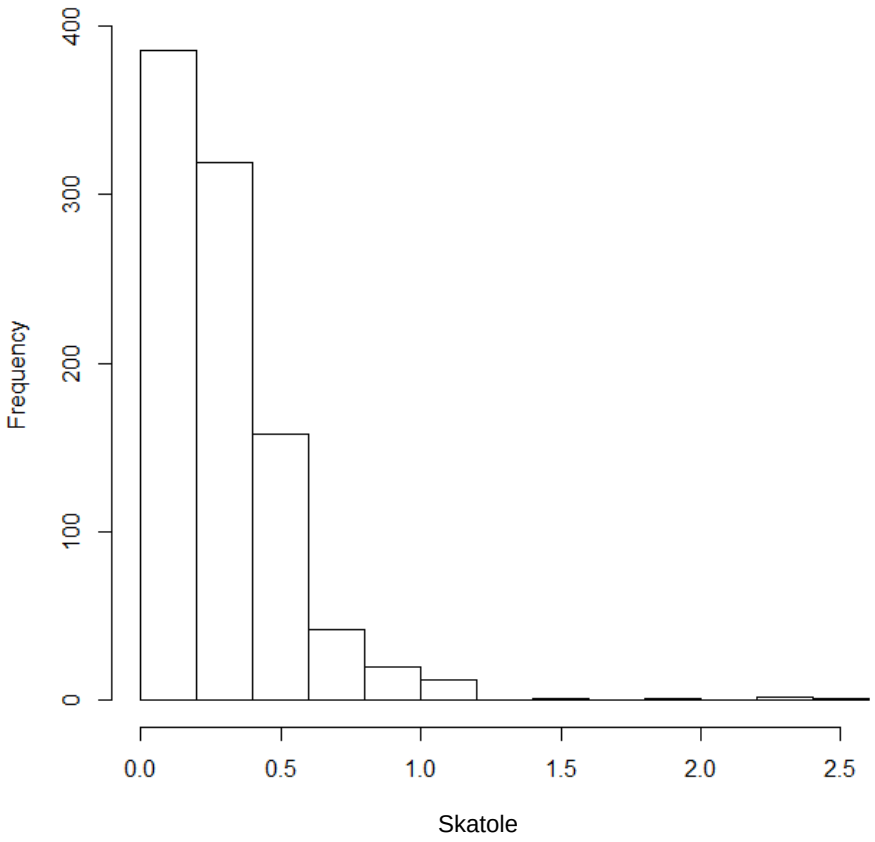


Figure 1. Distribution of skatole concentrations in untransformed form

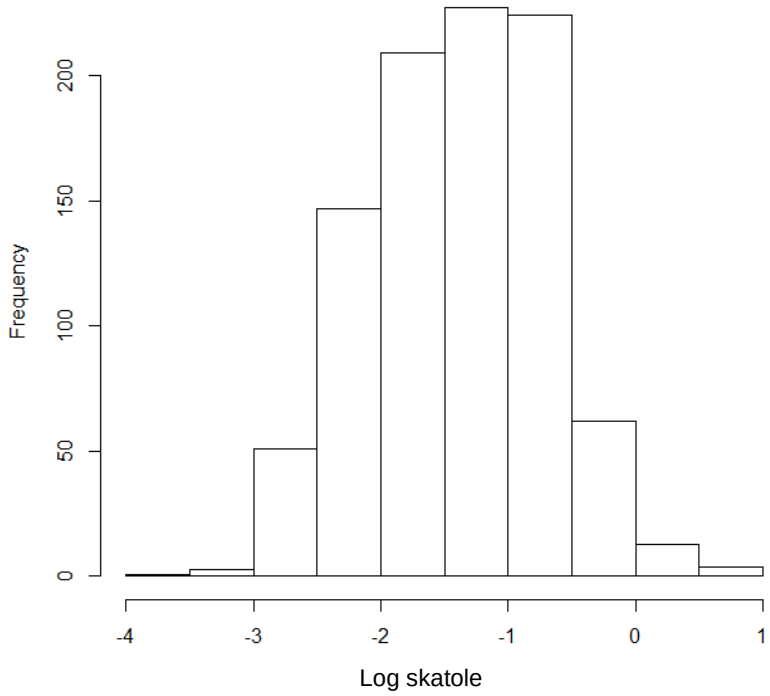


Figure 2. Distribution of skatole concentrations in log transformed form

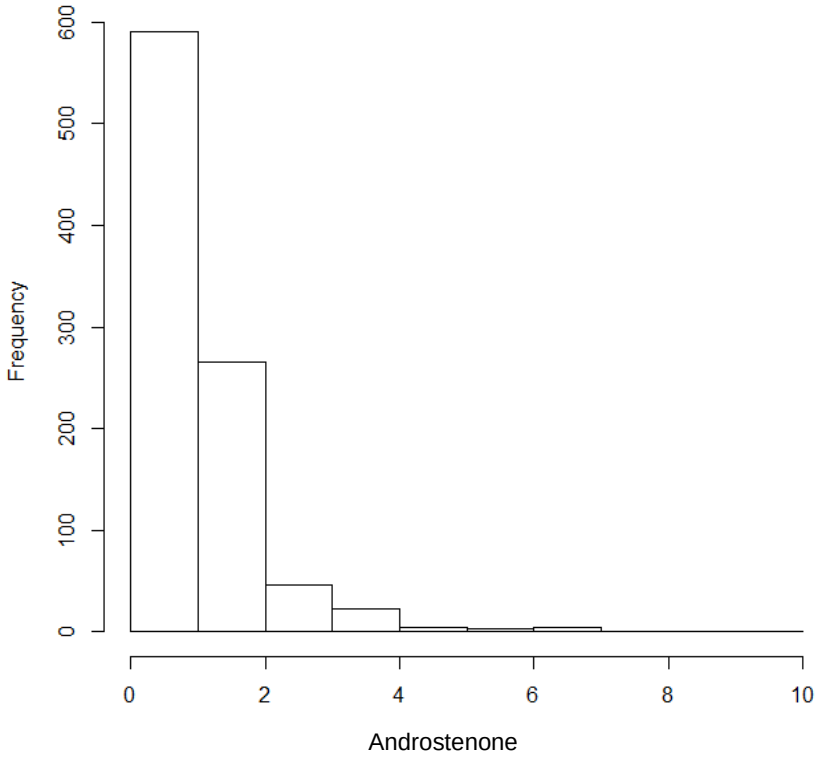


Figure 3. Distribution of androstenone concentrations in untransformed form

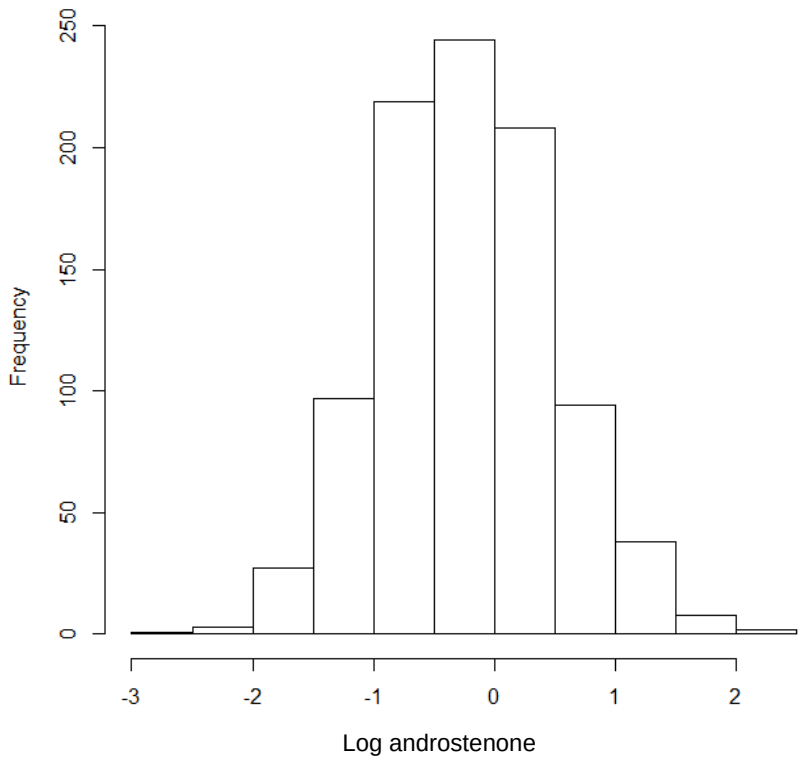


Figure 4. Distribution of androstenone concentrations in log transformed form

Distribution of androstenone concentrations are shown in *Figure 3.* and *Figure 4.* The concentration of androstenone in fat tissue (µg/g) was measured in all selected animals by time-resolved fluoro-immunoassay, as described by Tuomola et al., (1997), modified by using antiserum produced and characterized by Andresen (1974). Chemical analyses of skatole and androstenone were performed at Landbrug & Fødevarer, Denmark and the Norwegian School of Veterinary Science, respectively.

Data available

Information was collected on each animal including: sire, dam, age at slaughter, cold carcass weight, meat percentage, and the farm of rearing. The average age of selected animals at slaughter was 161.3 days while the average cold carcass weight was 77.34 kg. Average meat percentage was 60.13% determined by the standard Danish classification system in slaughterhouses.

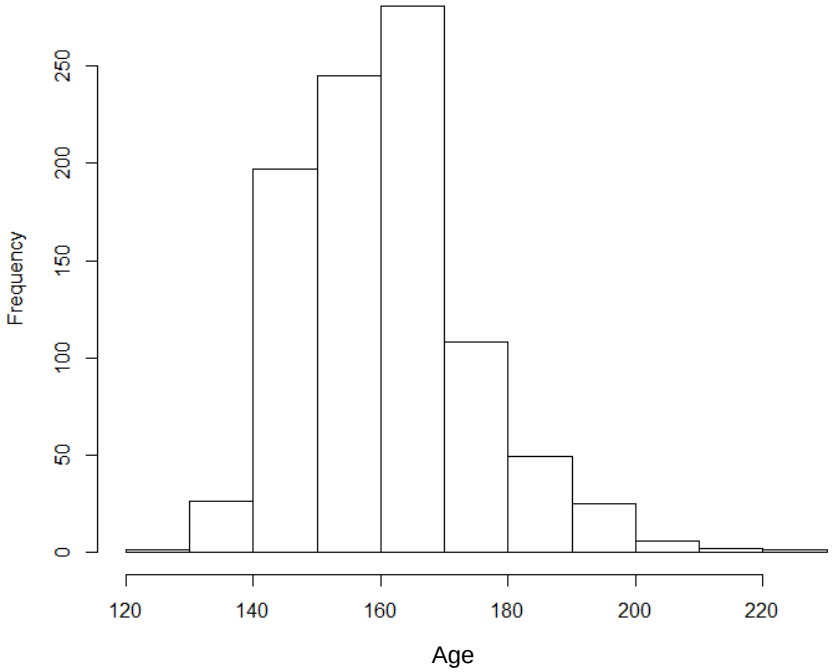


Figure 5. Distribution of age in days of the animals used

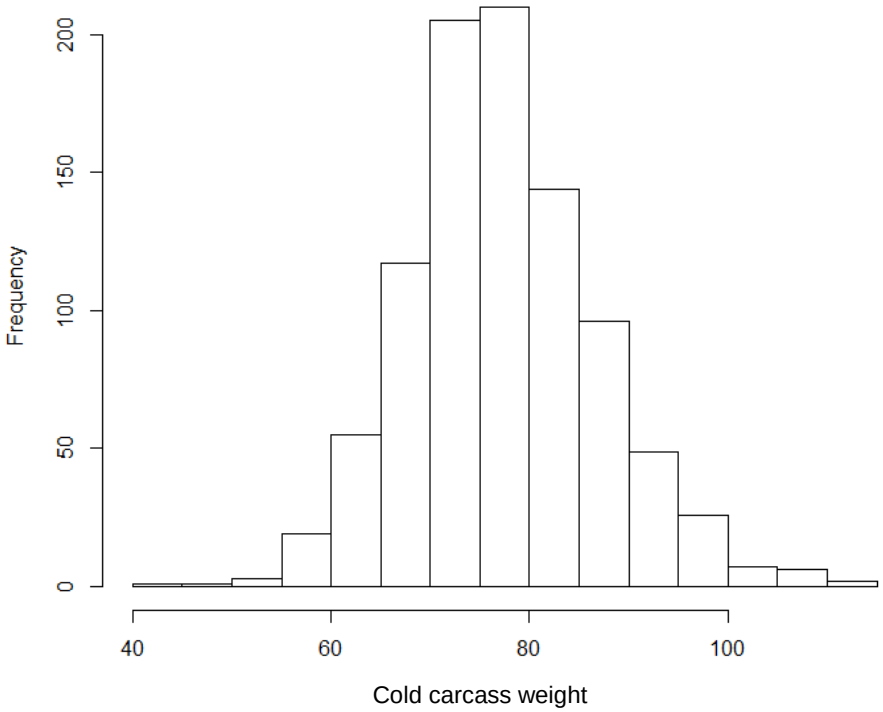


Figure 6. Distribution of cold carcass weight of the animals used

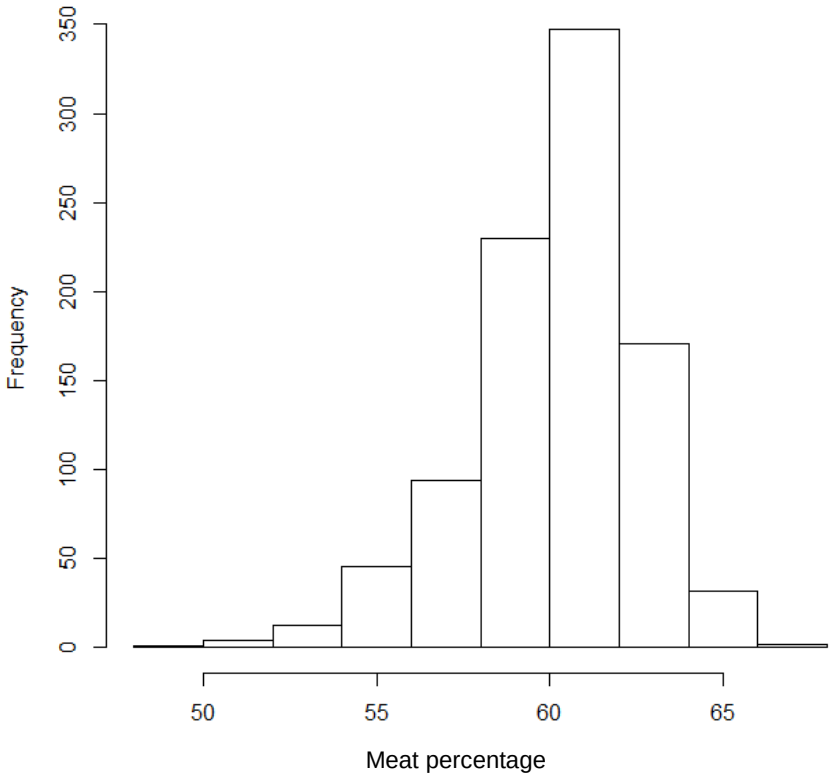


Figure 7. Distribution of meat percentage of the animals used

Genotyping and quality control

The 941 Danish Landrace boars were genotyped for 62,163 Single Nucleotide Polymorphisms (SNP) using the Illumina SNP60 porcine beadchip (Ramos et al., 2009). Quality control removed SNP loci with minor allele frequencies (MAF) < 0.01, call rate < 0.95, and those with extreme departure from Hardy-Weinberg equilibrium assessed using a FDR = 0.01. These criteria removed 13,795, 3,217 and 678 SNP, respectively. Three animals were excluded because of abnormally high autosomal heterozygosity. Therefore 938 animals with data on 42,916 SNPs (69%) remained in the analyses after quality control.

Table 6. Descriptive summary for genotypic data before quality control

Minor allele frequency distribution					
Class	X≤0.01	0.01<X≤0.05	0.05<X≤0.1	0.1<X≤0.2	X>0.2
Number	17749.0	4306.0	4212.0	8838.0	27058.0
Proportion	0.286	0.069	0.068	0.142	0.435

Cumulative distribution for SNPs at different significance thresholds for HWE					
Class	X≤10^{-4}	X≤0.001	≤0.01	XX≤0.05	all X
Number	212.0	461.0	1693.0	4453.0	62163
Proportion	0.003	0.007	0.027	0.072	1

Mean heterozygosity for a SNP	0.241
Standard deviation of the mean heterozygosity for a SNP	0.196
Mean heterozygosity for an individual	0.254
Standard deviation of mean heterozygosity for an individual	0.015

Table 7. Descriptive summary for genotypic data after quality control

Minor allele frequency distribution					
Class	X≤0.01	0.01<X≤0.05	0.05<X≤0.1	0.1<X≤0.2	X>0.2
Number	0	4214.0	4132.0	8565.0	26005.0
Proportion	0	0.098	0.096	0.2	0.606

Cumulative distribution for SNPs at different significance thresholds for HWE					
Class	X≤10^{-4}	X≤0.001	X≤0.01	XX≤0.05	all X
Number	0	0	987.0	3766.0	42916
Proportion	0	0	0.023	0.088	1

Mean heterozygosity for a SNP	
Standard deviation of the mean heterozygosity for a SNP	0.336
Mean heterozygosity for an individual	0.147
Standard deviation of mean heterozygosity for an individual	0.340

In addition, the litter mate design confirmed the expected population stratification due to the presence of closely related individuals. A clustering model was computed with the mclust function in R software 2.10 and multidimensional scaling (mds) was performed resulting in individuals being grouped into 3 clusters (Rowe et al., 2014) which separated some sire families. However there was no structural confounding observed between these clusters and the high and low skatole concentration groups because of the procedure for sampling animals for genotypes. This was confirmed in preliminary analyses by fitting the clusters as an independent factor in a linear model and no significant effect was observed.

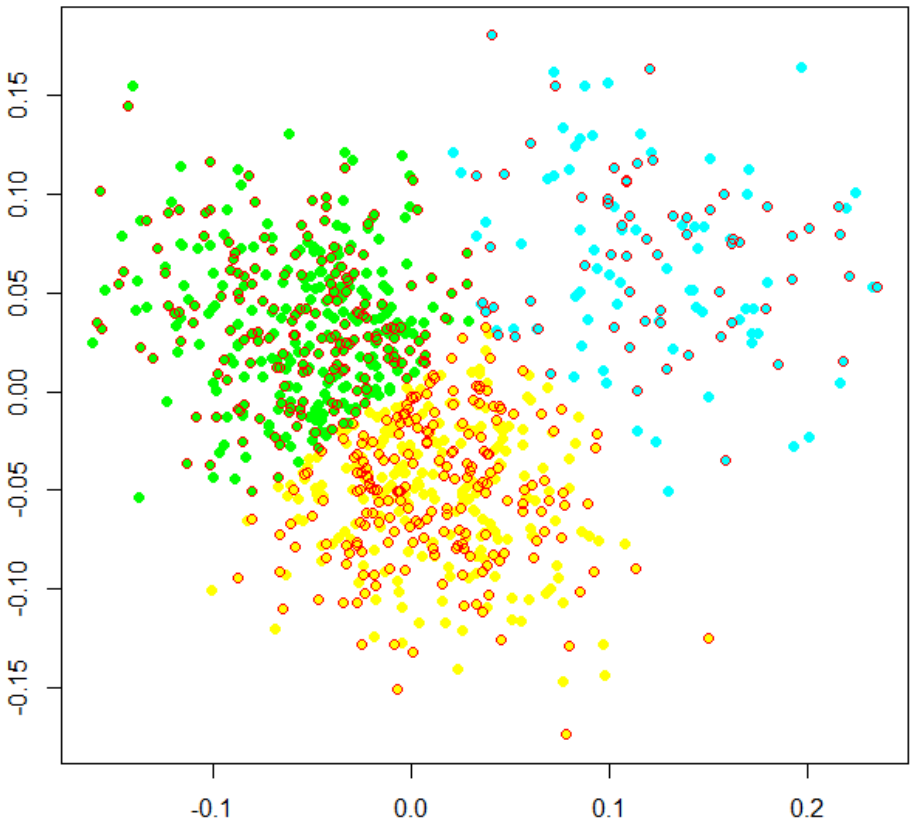


Figure 8. Plot of the three clusters (green, yellow and blue dots represent three detected clusters, while rounded red represents cases, respectively) using co-ordinates from multi-dimensional scaling

Methods of analyses

Phenotypic values for both traits were corrected for farm as a fixed effect and age as a covariate prior to genetic analysis. Meat percentage and cold carcass weight were not used as covariates as they could be confounded with genes that affect boar taint. The log-transformation was applied for skatole and androstenone phenotypic values in order to more closely approximate normal distributions. Six different models, GBLUP and five Bayesian variants, were fitted to both androstenone and skatole, as described below.

GBLUP. A mixed linear model was fitted as follows:

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{u} + \mathbf{e} \qquad (1)$$

where **y** is a vector of phenotypes of the trait; μ is the mean and **1** is vector of ones; **u** is a vector of random additive genetic effects assumed to be distributed MVN(0, $\sigma_g^2\mathbf{G}$) where **G** is a relationship matrix computed from the SNP information and constructed following Amin et.al. (2007) and σ_g^2 is the associated variance; and **e** is the vector of residuals assumed to be distributed MVN (0, $\sigma_e^2\mathbf{I}$) where **I** is the identity matrix. Amin et al. (2007) calculate **G** by:

$$g_{ij} = n^{-1}\sum_{k=1}^{n}(x_{ik} - 2p_k)(x_{jk} - 2p_k) / [2p_k(1-p_k)] \qquad (2)$$

$$g_{ii} = 1 + n^{-1}\sum_{k=1}^{n}(H_{E,k} - H_{ik}) / H_{E,k} \qquad (3)$$

where g_{ij} is the genomic relationship between animals *i* and *j*; x_{ik} is the genotype of the i^{th} individual at the k^{th} SNP when coded as 0, 1 and 2, for the reference allele homozygote, the heterozygote and alternative homozygote, respectively; p_k is the frequency of the reference

allele, n is the number of SNPs used for estimating relationships, $H_{E,k}$ is the expected heterozygosity at locus k and H_{ik} is the observed heterozygosity in animal i at locus k. This model was fitted using ASReml 3.0 (Gilmour *et.al.* 2000).

Bayesian regression methods. The linear model fitted for these methods was the following:

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{e} \qquad (4)$$

where, **y** is the vector of phenotypes; μ is overall mean for the trait and **1** is vector of ones; **Z** is the matrix of genotypes where z_{ik} is the number of alternative alleles for individual i at SNP locus k; and **β** is a vector of regression coefficients where β_k is the coefficient for SNP locus k; and **e** is the vector of residuals assumed to be distributed MVN (0, $\sigma_e^2\mathbf{I}$). The β_k are assumed to be independent random variables drawn from prior distributions which differ between the five Bayesian models.

The five models and their associated priors are as follows:

(1) **Bayes A**: The prior distribution for β_k is a scaled Student's *t*-distribution with two parameters scale, λ and shape υ.

(2) **Bayes B:** As Bayes A, but where only a fraction π of SNPs have effects from the scaled Student's *t*-distribution (with parameters scale λ and shape υ) with the remaining (1- π) have a zero effect.

(3) **Bayes C:** Similar to Bayes B but with non-zero effects assumed to be normally distributed with variance σ_s^2 instead of the scaled Student's *t*-distribution, and with mixing parameter π.

(4) **Bayes SSVS:** Similar to Bayes C but with effects coming from a mixture distribution of two normal distributions one with variance σ_s^2 and the other with variance σ_s^2/10000, and mixing parameter π (see Verbyla *et al.* 2009).

(5) **Bayesian Lasso:** Similar to Bayes A, but a Laplace distribution with scale parameter λ is replaces the scaled Student's *t*-distribution.

Frequently, the different parameters defining the prior distributions of β_k have been assumed as hyper-parameters and fixed in the analysis to a value pre-set by the researcher (e.g. Meuwissen et al. 2001, Hayes et al. 2009). Here, these parameters were included in the analysis and estimated from the data, with the exception of π as the low heritability of skatole made the analysis prone to convergence problems when using Bayes C, where it was fixed to be 0.1, but preliminary analysis showed that the results were similar over a range of small values for π. For all the other parameters defining the distributions of SNP effects, a bounded flat prior was assumed. The scale parameter λ (included in Bayes A, Bayes B and Bayesian Lasso), the variance parameter σ_s^2 (included in Bayes C and Bayes SSVS), and the residual variance σ_e^2 were all bounded between 0 and a very large positive number so that any influence of the prior on the estimated genetic variance was negligible. The shape parameter υ in Bayes A and Bayes B were bounded between 0.5 and 8.

The implementation of the Bayesian regression method was carried out using Gibbs sampling. For each of the analysis carried out here the first 50000 cycles of the Monte Carlo Markov Chain were discarded as a burn-in period. Results were calculated from a minimum of 20000 subsequent realisations where consecutive realization was separated by 50 cycles. The whole chain therefore consisted of 1,050,000 cycles.

Cross validation and comparisons between the methods

A 5-fold cross-validation was carried out to compare the accuracy of GBLUP and the five Bayesian methods: Bayes A, Bayes B, Bayes C, Bayes SSVS and Bayesian Lasso to predict the unobserved phenotypes. The division of the full dataset preserved sib pairs but was otherwise randomly separated into five cross-validation sets resulting in training sets of ~751 animals and validation sets ~187 animals. Each training set had a size of approximately a 4/5 of the whole dataset with phenotypes and each animal appeared in precisely one validation set. For each training set, GBLUP and Bayesian methods were used to estimate GEBV and heritabilities. Accuracy (r) to predict the phenotype was calculated as the correlation between the GEBV and the phenotypes of validation animals and the overall values of accuracies were calculated as average over the five validation sets. Principal component analyses for both traits were performed in order to show the relative relationship between all the methods investigated.

Comparisons with QTL

The difference between genomic predictions using all SNP and an approach utilising only SNP identified from GWAS was assessed by calculating the predictive accuracy of all SNP identified as statistically significant ($P<0.05$) genome-wide from the same dataset (Rowe et al., 2014). These SNP were H3GA00016037 on SSC5 for androstenone concentrations and SIRI0000194 on SSC14 for skatole concentrations. This was done using the 5 cross-validations sets with the phenotype of each set being predicted using estimates of the magnitude of the QTL-effect derived by estimating allelic substitution effects by fitting SNP genotypes (coded as 0, 1 and 2) to the remaining data.

Genomic BLUP with dominance effects included

For the purpose of GBLUP analyses with dominance effects included, the statistical model was extended. Partitioning the total genetic variance into additive and dominance genetic variance required construction of two separate genomic relationship matrices, G and D, for additive and dominance, respectively. These matrices can be calculated from the genome-wide SNPs and they describe the relationships between genotyped individuals. Construction of additive genomic relationship matrix was constructed in a same way as used in previous GBLUP analyses, described above. Dominance genomic matrix was constructed based on the following principles. Let A_{1k} and A_{2k} be 2 alleles at the k^{th} marker locus and p_k be the frequency of A_{2k}. The dominance genotype values for SNP (k) are 0 if the animal is any of the two homozygotes, $A_1 A_1$, $A_2 A_2$ respectively, and 1 if the animal is heterozygote $A_1 A_2$. Dominance deviation of an individual is calculated as follows:

$$\boldsymbol{M}_{\boldsymbol{d}ij} = \begin{cases} -2p_k^2\ (A_1A_1) \\ 2p\ (1-p_k)\ (A_1A_2) \\ -2\ (1-p_k)^2 (A_2A_2) \end{cases} \qquad (5)$$

$\mathbf{M}_{\mathbf{d}ij}$ is the n x N_m matrix where n is a number of individuals.

Therefore, using M_d the dominance genomic relationship matrix can be calculated as:

$$\mathbf{D} = \frac{1}{n} \sum_{k=1}^{n} \frac{M_d M_d'}{\{2p_k(1-p_k)\}^2} \qquad (6)$$

In the models the dominance genetic effects were assumed to follow a multivariate normal distributions: MVN(0, $\sigma^2_d \mathbf{D}$).

Mixed model using whole genomic and regional genomic approach

Mixed model equations were constructed for estimation of variance components using whole genome approach. The analyses were based on fitting random additive (GBLUP-A) and dominance genetic effects (GBLUP-D) of all markers across the genome. Also, the model with both additive and dominance genetic effects fitted was named GBLUP-AD. In order to account for the possible maternal effect of the dam on the traits, dam information was included in a model as an additional independent random effect. In traditional pedigree analyses the major source of information on dominance variance is from full-sib families and this is often confounded with maternal effects. The equation was as follows:

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{u} + \mathbf{d} + \mathbf{Zm} + \mathbf{e} \qquad (7)$$

where **y** is a vector of phenotypes of the trait; μ is the mean and **1** is vector of ones; **u** is a vector of random additive genetic effects assumed to be distributed MVN(0, σ^2_g **G**) where **G** is a relationship matrix computed from the SNP information and constructed as described above, and σ^2_g is the associated variance; **d** is a vector of random dominance genetic effects assumed to be distributed MVN(0, σ^2_d **D**) where **D** is a relationship matrix computed from the SNP information and constructed as described above, and σ^2_d is the associated variance; **m** is a vector of random dam effects assumed distributed MVN(0, σ^2_m **I**) where σ^2_m is the associated variance is the variance matrix for the dam effects and **I** is the identity matrix and **Z** is the design matrix linking dams to offspring; and **e** is the vector of residuals assumed to be distributed MVN (0, σ^2_e **I**). Variance components were estimated with average information restricted maximum likelihood (REML).

Besides the whole genome approach, a regional approach was also performed for further analysis of variance components. Within this analysis, the genome was divided on 18 autosomal chromosomes and regional chromosomal genomic relationship matrices were created separately for each of the 18 autosomal chromosomes in order to estimate random additive and dominance genetic effects attributable to each autosomal chromosome. Furthermore, for the analysis of single chromosomal variance components, additional genomic relationship matrices were created using SNPs on the remaining chromosomes to account for the genetic variation complementary to the chromosome. This strategy was used in order to test the likelihood ratio of proposed models (explained in the next paragraph). Therefore, three models were used to estimate regional chromosomal variance components for each autosomal chromosome. The equations were as follows:

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{u_c} + \mathbf{u_{-c}} + \mathbf{d_c} + \mathbf{d_{-c}} + \mathbf{Zm} + \mathbf{e} \qquad (8)$$

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{u_c} + \mathbf{u_{-c}} + \mathbf{d_{-c}} + \mathbf{Zm} + \mathbf{e} \qquad (9)$$

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{u_{-c}} + \mathbf{d_c} + \mathbf{d_{-c}} + \mathbf{Zm} + \mathbf{e} \qquad (10)$$

where **y** is a vector of phenotypes of the trait; μ is the mean and **1** is vector of ones; $\mathbf{u_c}$ is a vector of random additive genetic effects of each chromosome assumed to be distributed MVN $(0, \sigma^2_{g(c)}\,\mathbf{G_c})$ where $\mathbf{G_c}$ is a relationship matrix computed from the SNP information of each chromosome, and $\sigma^2_{g(c)}$ is the associated variance; $\mathbf{u_{-c}}$ is a vector of random additive genetic effects of remaining chromosomes assumed to be distributed MVN $(0, \sigma^2_{g(-c)}\,\mathbf{G_{-c}})$ where $\mathbf{G_{-c}}$ is a relationship matrix computed from the SNP information of remaining chromosomes, and $\sigma^2_{g(-c)}$ is the associated variance; $\mathbf{d_c}$ is a vector of random dominance

genetic effects of each chromosome assumed to be distributed MVN (0, $\sigma^2_{d(c)}$ $\mathbf{D_c}$) where $\mathbf{D_c}$ is a relationship matrix computed from the SNP information of each chromosome, and $\sigma^2_{d(c)}$ is the associated variance; $\mathbf{d_{-c}}$ is a vector of random dominance genetic effects of remaining chromosomes assumed to be distributed MVN (0, $\sigma^2_{d(-c)}$ $\mathbf{D_{-c}}$) where $\mathbf{D_{-c}}$ is a relationship matrix computed from the SNP information of remaining chromosomes and constructed as described above, and $\sigma^2_{d(-c)}$ is the associated variance; **m** is a vector of random dam genetic effects where **Z** is incident matrix for the dam effects and **e** is the vector of residuals assumed to be distributed MVN (0, σ^2_e **I**) where **I** is the identity matrix.

Model validation

In order to test for the significance of proposed models against the null hypothesis the likelihood ratio (LRT) test statistic was calculated LRT = $-2\ln(L_0/L_1)$, where L_0 and L_1 stands for the obtained likelihood values under the null hypothesis (H_0) or proposed model (H_1). When REML is used to estimate genetic and environmental components of variance the asymptotic distribution of the LRT is a mixture of χ^2 distributions with different degrees of freedom (Visscher, 2006). Therefore, for a 5% significance level for a test of a single component model with H_o: $\sigma^2 = 0$, the value for 10% significance level of χ_1^2 is considered which is 2.71.

To test the significance of estimated genetic variance components in the whole genome approach, LRT was used to compare the null hypothesis of the model fitting either additive (GBLUP-A) or dominance (GBLUP-D) variance against the model with both additive and dominance (GBLUP-AD) variance estimated.

For the regional approach, an additional strategy was used for testing the goodness of fit between two models. The likelihood ratio was calculated (LRTadd) whether fitting

individual chromosome (equation (9)) could explain additional proportion of additive genetic variance against the null hypothesis of the model with the rest of the genome (equation (10)). The same strategy was used to test (LRTdom) the proportion of dominance genetic variance (equation (8)) of each chromosome against the null hypothesis of the model with the rest of the genome (equation (9)).

Heritability estimates

From the models used, four variance components were estimated: additive genetic variance, dominance genetic variance, dam maternal variance and environmental variance.

Let the σ_p^2 be the total phenotypic variance partitioned as follows:

$$\sigma_p^2 = \sigma_a^2 + \sigma_d^2 + \sigma_m^2 + \sigma_e^2$$

where σ_a^2 is additive genetic variance, σ_d^2 is dominance genetic variance, σ_m^2 is dam maternal variance and σ_e^2 is environmental variance.

Based on those estimates we calculated three heritability estimates as follows:

Narrow sense heritability, $h_a^2 = \sigma_a^2 / \sigma_p^2$; broad sense heritability $H^2 = (\sigma_a^2 + \sigma_d^2) / \sigma_p^2$. Furthermore, to further show the proportion of dam maternal variance in total variation, ratio to phenotypic variance was calculated as $m^2 = \sigma_m^2 / \sigma_p^2$, respectively.

5. RESULTS

Androstenone. Table 8 shows the accuracies (average correlation between the GEBV and phenotypes across the validation sets) obtained by the different methodologies. The range of accuracies for predicting phenotype was narrow for androstenone, only ranging between 0.291 (Lasso) and 0.310 (Bayes B), 6% of the mean accuracy, and with no clear difference between GBLUP and Bayesian methodologies. The estimated h^2 were also narrow ranging from 0.276 (Bayesian Lasso) to 0.307 (GBLUP). GBLUP also had the lowest σ^2_e, which is the most objective component for comparison since its magnitude does not depend on scaling assumptions, but the range of estimates was only 4% of their mean. Scaling all the accuracies of predicting phenotypes by the square root of the average h^2 indicated accuracies of predicting the breeding value of ~ 0.56.

Skatole. The heritabilities and accuracies calculated as correlations between the estimated genomic breeding values (GEBV) and phenotypes of the validation animals from different methodologies are shown in table 9. Compared to androstenone, the range of accuracies for predicting skatole fat concentrations was wider, between 0.214 (GBLUP) and 0.266 (Bayes SSVS and C), corresponding to 21% of the mean over all methods, with GBLUP appearing to be a low outlier. In contrast the range in estimates of σ_e^2 was very similar to androstenone corresponding to 4% of the mean estimate over methods. The estimated heritability was highest with the Bayes C method (0.106) and lowest with GBLUP (0.051). Using the average of the estimates, the accuracy of predicting the breeding value was 0.88.

Comparison of methods

Figures 9 and 10 shows the relationships between individual SNP effects across methods. The plot confirms the strong similarity between Bayes B and Bayes SSVS, and, in turn, their similarity with Bayes A. All three methods have the assumption that large SNP effects follow an inverse chi-squared distribution. Bayes C shows a narrower range of values compared to these, as might be expected from the regularisation properties of these distributions. The SNP effects for Bayesian Lasso had the lowest variance of all methods.

For skatole, where a single, strong QTL is present (Rowe et al., 2014), the best accuracy was obtained by Bayes SSVS, followed by Bayes C. Bayesian Lasso performed similarly for both traits, achieving the lowest accuracy as well as lowest proportion of genetic variance captured.

In order to further demonstrate relative relationships between the used methodologies, principal component analysis was performed on GEBV-s and the obtained results are presented in figures 11 and 12. As expected, the scatter plot indicates greater similarity amongst the methodologies for the estimation of GEBV-s for androstenone, than for skatole. This perspective of the different methods was confirmed with the PCA analysis of the genomic EBV (see Figure 2). For both androstenone and skatole Bayes A, B and SSVS tended to cluster together, and Bayesian Lasso clustered with GBLUP but the differences were small.

Comparison with QTL

For androstenone, the accuracy of predicting phenotypes from the single significant SNP was 0.15, notably lower than the other genomic predictions using all SNP. For skatole concentration the accuracy in predicting phenotypes from the single genome wide significant SNP was 0.21, similar to GBLUP but lower than those obtained for Bayesian methodologies.

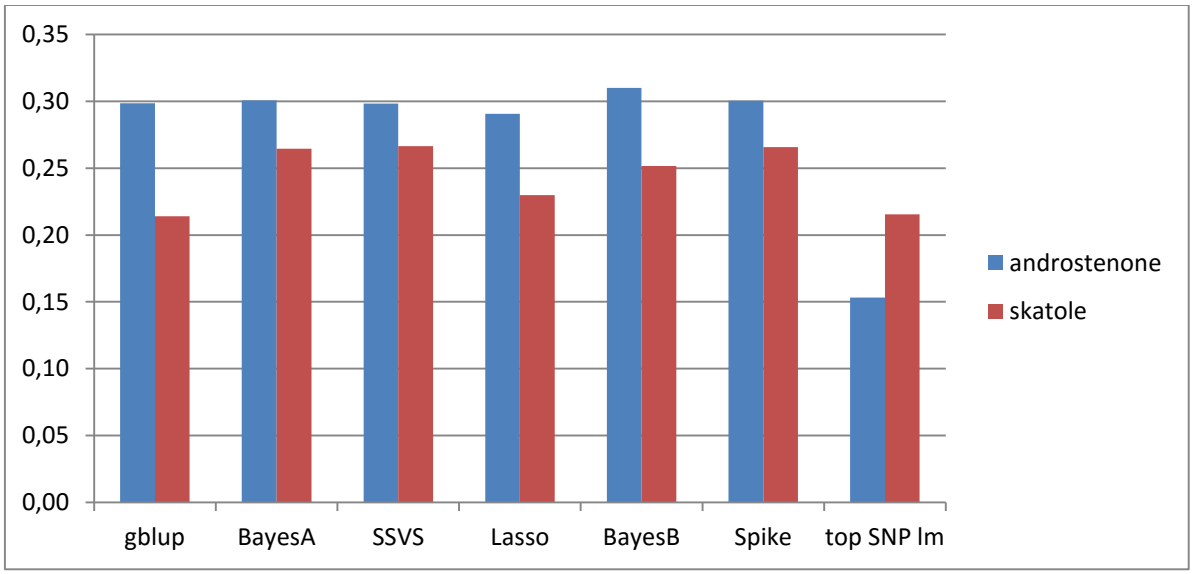


Figure 7. Accuracies (r) of predicting androstenone and skatole concentrations estimated with all methodologies used including single QTL

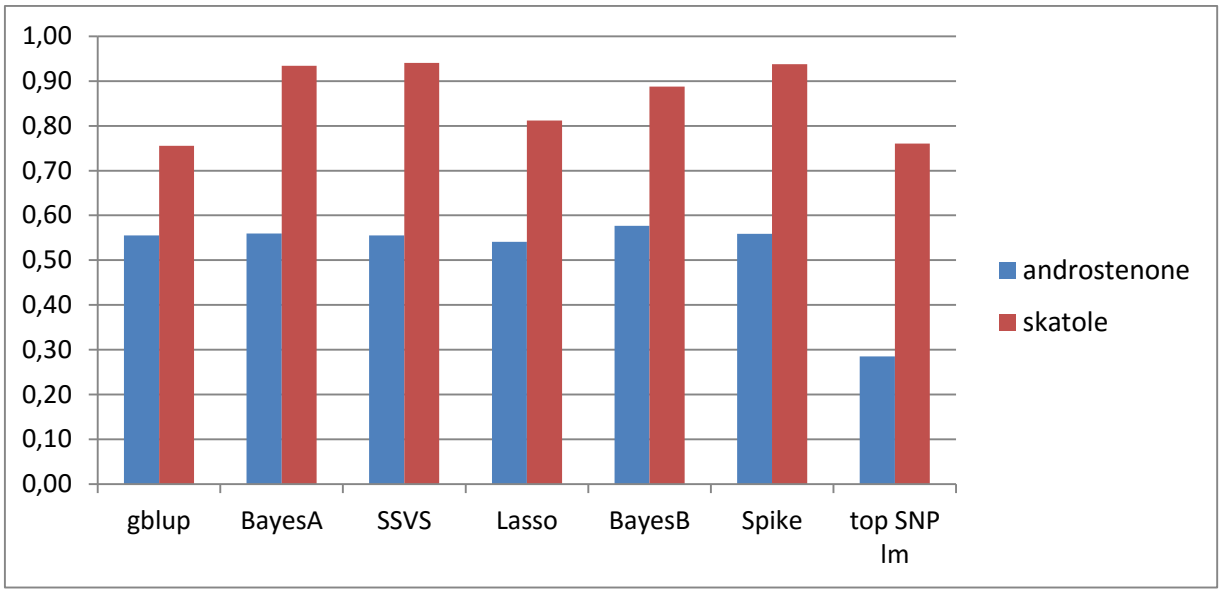


Figure 8. Accuracies (r) of predicting androstenone and skatole concentrations estimated with all methodologies used including single QTL, and scaled by the square root of the average h^2 over all methods.

Table 8. Genetic (σ^2_g) and residual (σ^2_e) variance components, heritabilities (h^2) and accuracies (r and r) for androstenone concentration (µg/g fat tissue) estimated by different methodologies*

Method	σ^2_g	σ^2_e	h^2	**r**	r^*
GBLUP	0.149	0.333	0.307	0.298	0.555
Bayes A	0.141	0.343	0.287	0.301	0.559
Bayes B	0.137	0.347	0.276	0.310	0.577
Bayes SSVS	0.143	0.343	0.281	0.299	0.555
Bayes C	0.149	0.337	0.299	0.300	0.559
Bayesian LASSO	0.137	0.346	0.284	0.291	0.541

r is the accuracy of predicting the phenotype calculated as the correlation between the estimated breeding value and phenotype; r* corresponds to the accuracy of the breeding value estimate, obtained by scaling r by the square root of the average h^2 over all methods

Table 9. Genetic (σ^2_g) and residual (σ^2_e) variance components, heritabilities (h^2) and accuracies (r and r*) for skatole concentration (μg/g fat tissue) estimated by different methodologies.

Method	**σ^2_g**	**σ^2_e**	**h^2**	**r**	**r***
GBLUP	0.014	0.466	0.051	0.214	0.755
Bayes A	0.037	0.446	0.094	0.265	0.934
Bayes B	0.030	0.452	0.074	0.252	0.888
Bayes SSVS	0.039	0.446	0.087	0.266	0.940
Bayes C	0.037	0.447	0.106	0.266	0.938
Bayesian LASSO	0.028	0.457	0.068	0.230	0.812

r is the accuracy of predicting the phenotype calculated as the correlation between the estimated breeding value and phenotype; r* corresponds to the accuracy of the breeding value estimate, obtained by scaling r by the square root of the average h^2 over all methods

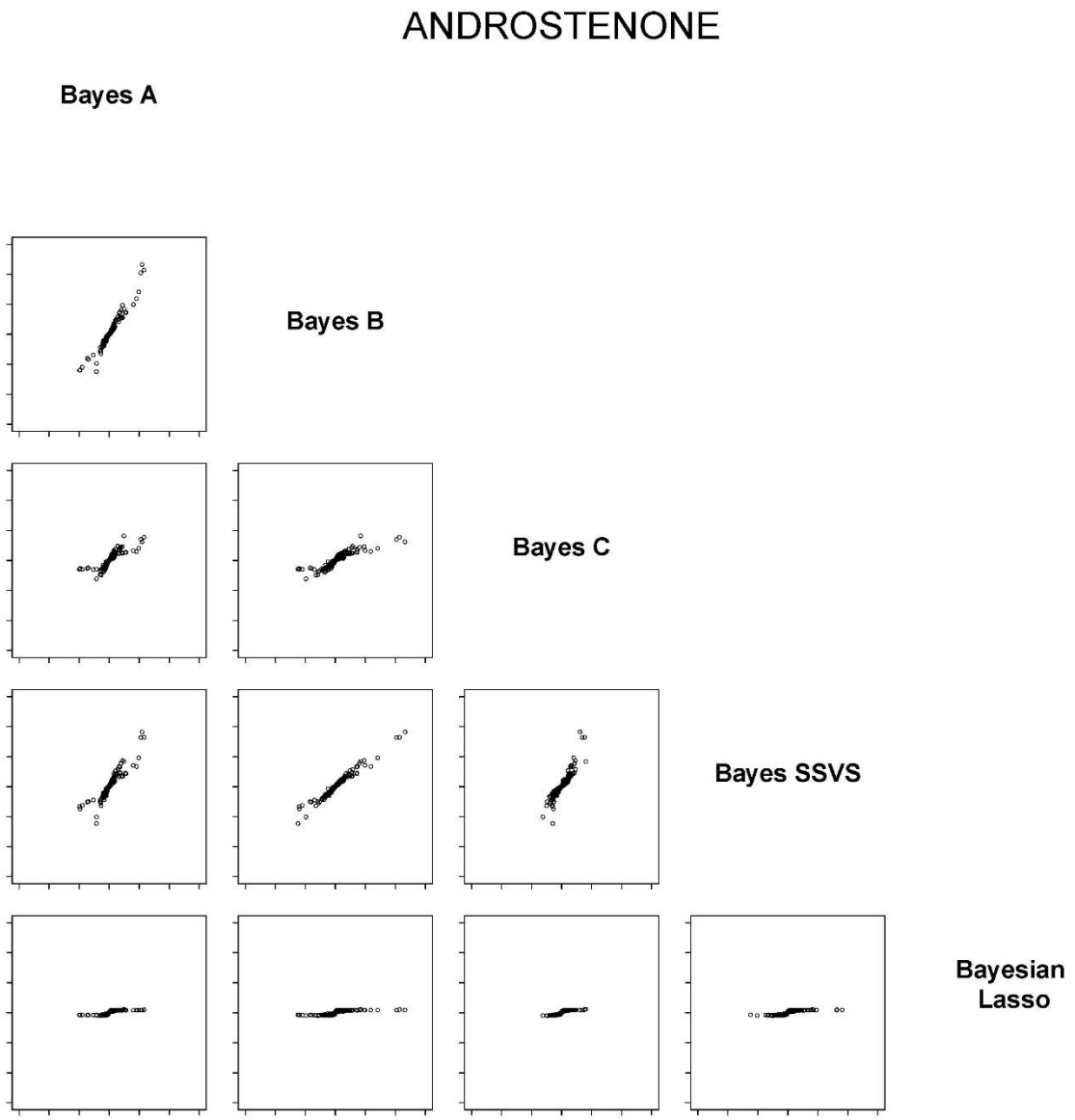


Figure 9. A comparison of estimated SNP effects, defined as the average value over realisations, obtained for five Bayesian methods for androstenone (measured as μg/g fat tissue). Coordinate length for both x and y axes ranges from -0.03 to 0.03.

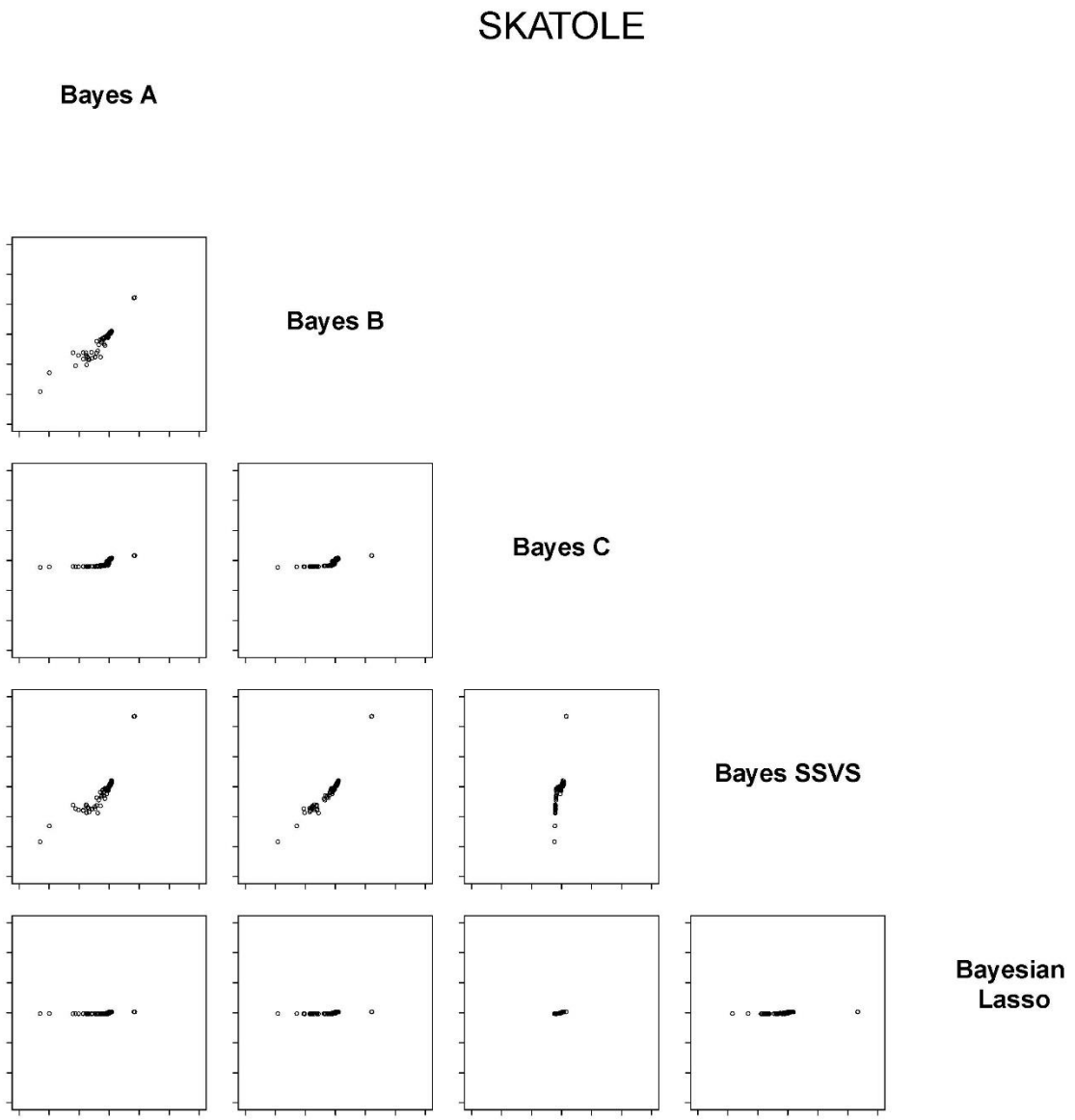


Figure 10. A comparison of estimated SNP effects, defined as the average value over realisations, obtained for five Bayesian methods for skatole (measured as μg/g fat tissue). Coordinate length for both x and y axes ranges from -0.03 to 0.03.

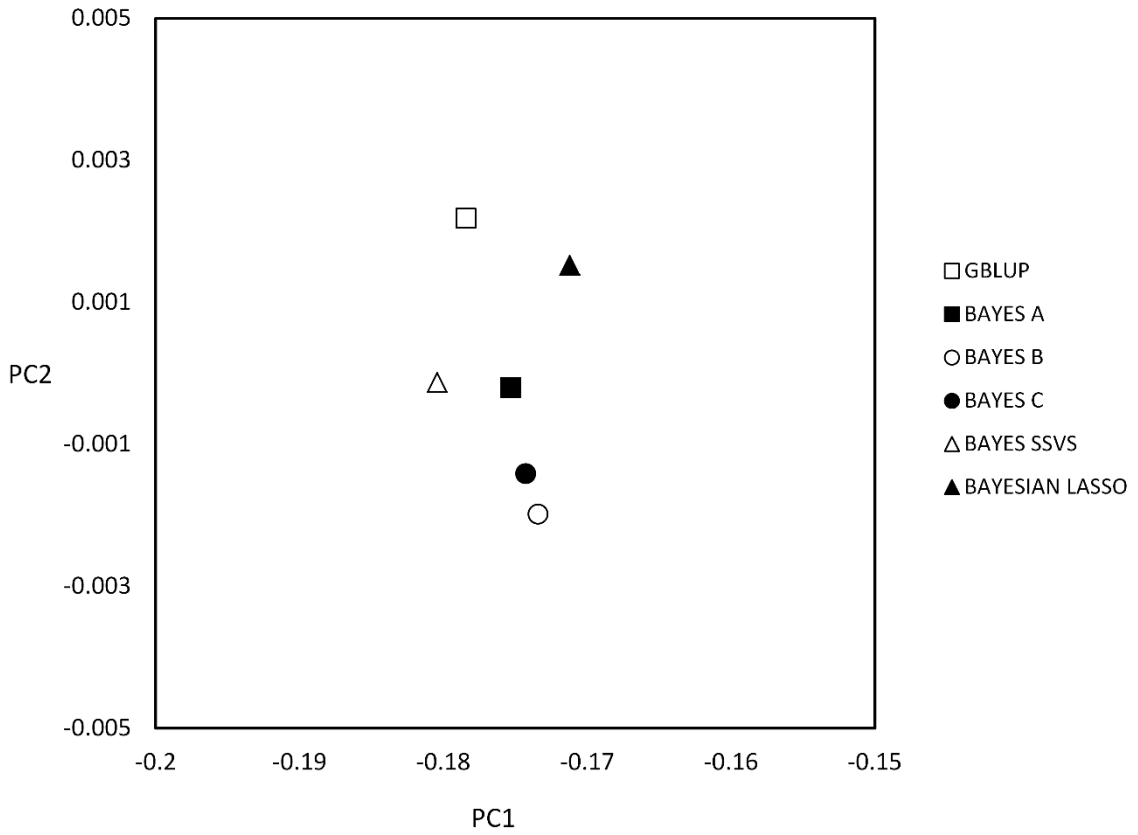


Figure 11. Scatterplot of the first two principal components (PC1 vs. PC2) on the GEBV for androstenone concentrations between all the methods. Each point represents different method as follows: GBLUP (□), Bayes A (■), Bayes B (○); Bayes C (●), Bayes SSVS (Δ), Bayesian Lasso (▲).

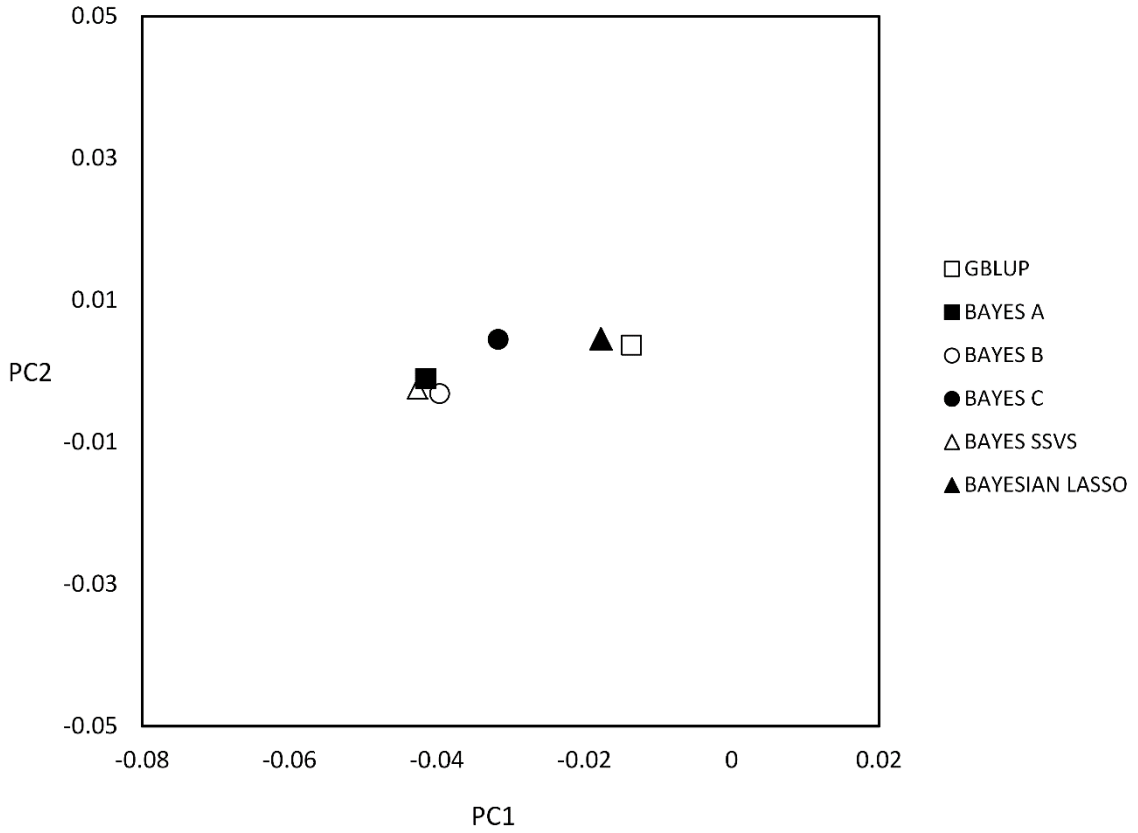


Figure 12. Scatterplot of the first two principal components (PC1 vs. PC2) on the GEBV for skatole concentrations between all the methods. Each point represents different method as follows: GBLUP (□), Bayes A (■), Bayes B (○); Bayes C (●), Bayes SSVS (Δ), Bayesian Lasso (▲).

Estimation of dominance effects

Table 10. Genetic (σ_a^2, σ_d^2, σ_m^2) and residual (σ_e^2) variance components, heritabilities and ratio of dam genetic variance to total variance for androstenone estimated by GBLUP under three models

Androstenone	σ_a^2	σ_d^2	σ_m^2	σ_e^2	σ_y^2	h^2	H^2	m^2	**LRT**
GBLUP-A	0,135		0,043	0,302	0,480	0,282		0,089	3,443
GBLUP-D		0,165	0,034	0,277	0,476		0,347	0,071	17,653
GBLUP-AD	0,108	0,063	0,035	0,274	0,479	0,225	0,357	0,072	

Table 11. Genetic (σ_a^2, σ_d^2, σ_m^2) and residual (σ_e^2) variance components, heritabilities and ratio of dam genetic variance to total variance for skatole estimated by GBLUP under three models

Skatole	σ_a^2	σ_d^2	σ_m^2	σ_e^2	σ_y^2	h^2	H^2	m^2	**LRT**
GBLUP-A	0,026		0,000	0,453	0,480	0,055		0,000	0,088
GBLUP-D		0,006	0,000	0,474	0,480		0,012	0,000	3,601
GBLUP-AD	0,026	0,001	0,000	0,453	0,480	0,054	0,056	0,000	

Table 12. Ratios of chromosomal additive ($\sigma_a{}^2$) and dominance ($\sigma_d{}^2$) genetic variances in total phenotypic variance and significance of the LRT for androstenone

Chromosome	Additive $\frac{\sigma^2_{a\,chr}}{\sigma^2_{a\,chr} + \sigma^2_{a\,(-chr)}}$	Dominance $\frac{\sigma^2_{d\,chr}}{\sigma^2_{d\,chr} + \sigma^2_{d\,(-chr)}}$	Significance *a*	Significance *d*
1	0.00	0.00	n.s.	n.s.
2	0.01	0.00	n.s.	n.s.
3	0.01	0.04	n.s.	n.s.
4	0.00	0.04	n.s.	n.s.
5	0.02	0.00	n.s.	n.s.
6	0.01	0.00	n.s.	n.s.
7	0.00	0.00	n.s.	n.s.
8	0.01	0.07	n.s.	n.s.
9	0.01	0.11	n.s.	n.s.
10	0.00	0.18	n.s.	n.s.
11	0.00	0.27	n.s.	*
12	0.10	0.00	n.s.	n.s.
13	0.00	0.00	n.s.	n.s.
14	0.00	0.03	n.s.	n.s.
15	0.16	0.00	n.s.	n.s.
16	0.00	0.09	n.s.	n.s.
17	0.00	0.03	n.s.	n.s.
18	0.00	0.00	n.s.	n.s.

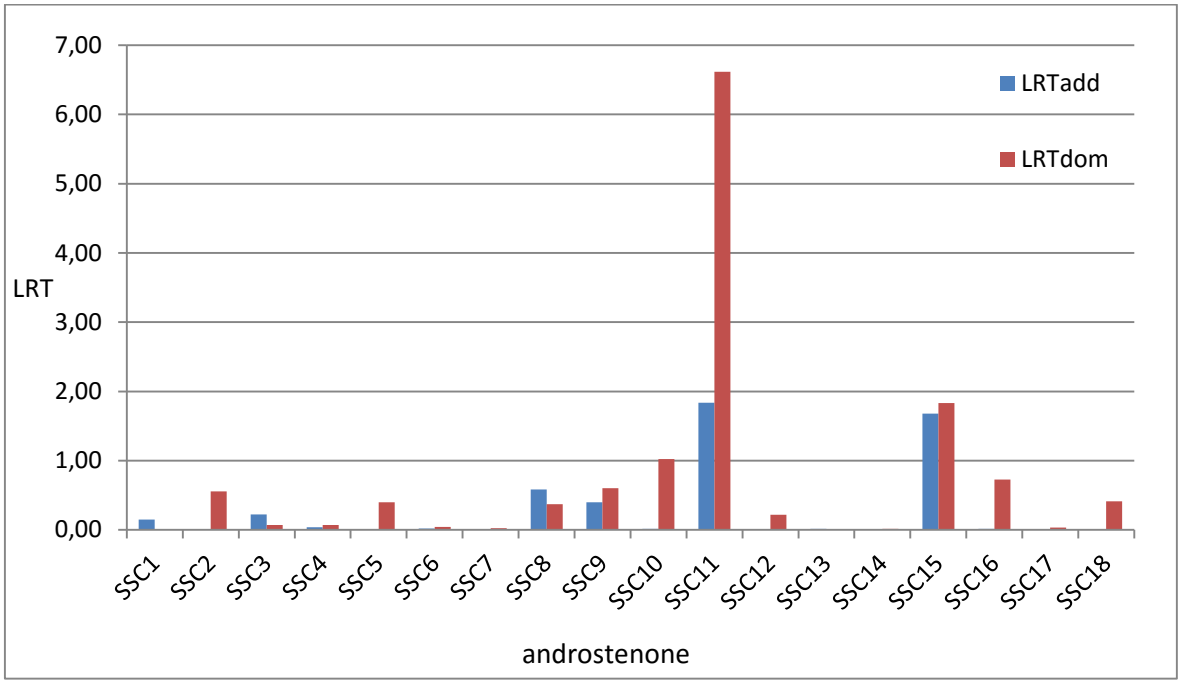


Figure 13. Likelihood ratio test (LRT) for significance of additive and dominance models against the null model for androstenone

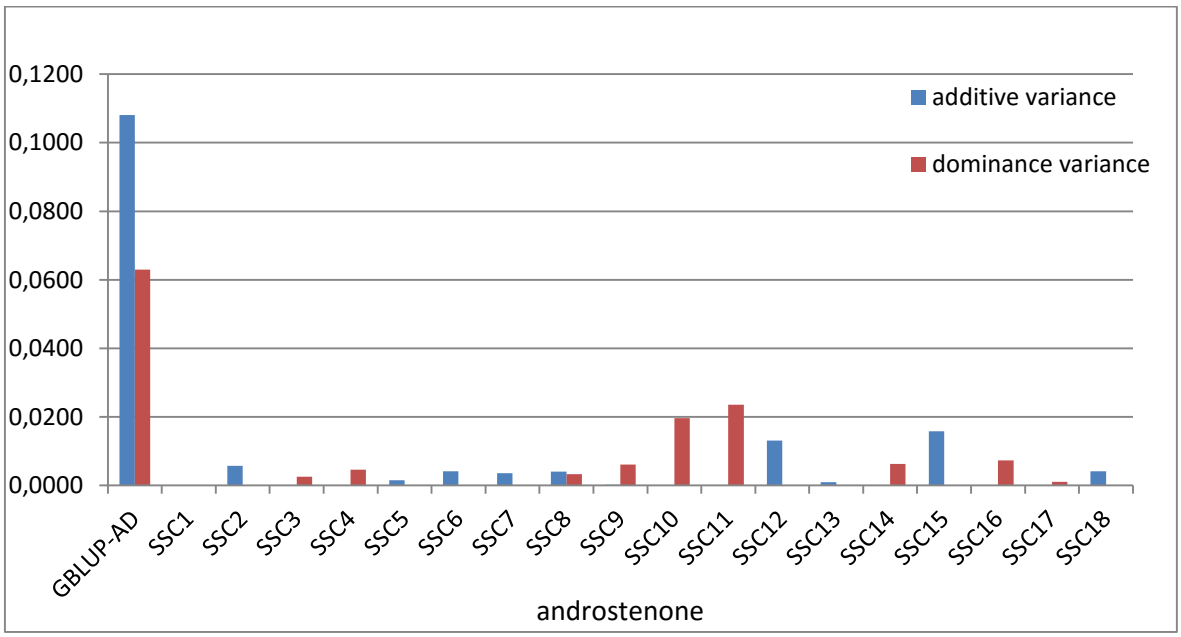


Figure 14. Genetic (σ_a^2 and σ_d^2) variance components for androstenone estimated by GBLUP-AD for whole genome and regional heritability approach

Table 13. Ratios of chromosomal additive ($\sigma_a{}^2$) and dominance ($\sigma_d{}^2$) genetic variances in total phenotypic variance and significance of the LRT for skatole

Chromosome	Additive $\frac{\sigma^2_{a\,chr}}{\sigma^2_{a\,chr} + \sigma^2_{a\,(-chr)}}$	Dominance $\frac{\sigma^2_{d\,chr}}{\sigma^2_{d\,chr} + \sigma^2_{d\,(-chr)}}$	Significance *a*	Significance *d*
1	0.01	0.38	n.s.	n.s.
2	0.06	0.87	n.s.	n.s.
3	0.01	0.27	n.s.	n.s.
4	0.01	0.50	n.s.	n.s.
5	0.01	0.19	n.s.	n.s.
6	0.03	0.88	n.s.	n.s.
7	0.03	0.25	n.s.	n.s.
8	0.01	0.40	n.s.	n.s.
9	0.02	0.96	n.s.	*
10	0.01	0.53	n.s.	n.s.
11	0.02	0.29	n.s.	n.s.
12	0.02	0.21	n.s.	n.s.
13	0.01	0.53	n.s.	n.s.
14	0.22	0.49	n.s.	n.s.
15	0.20	0.48	n.s.	n.s.
16	0.07	0.87	n.s.	n.s.
17	0.01	0.21	n.s.	n.s.
18	0.01	0.32	n.s.	n.s.

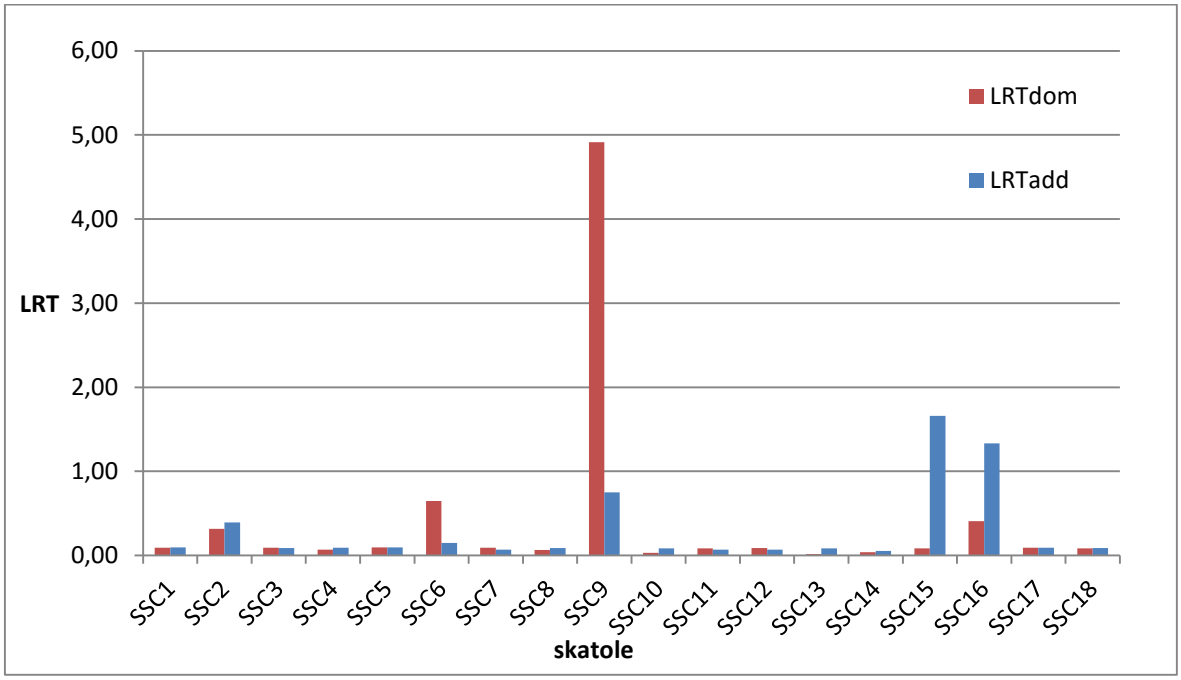


Figure 15. Likelihood ratio test (LRT) for significance of additive and dominance models against the null model for skatole

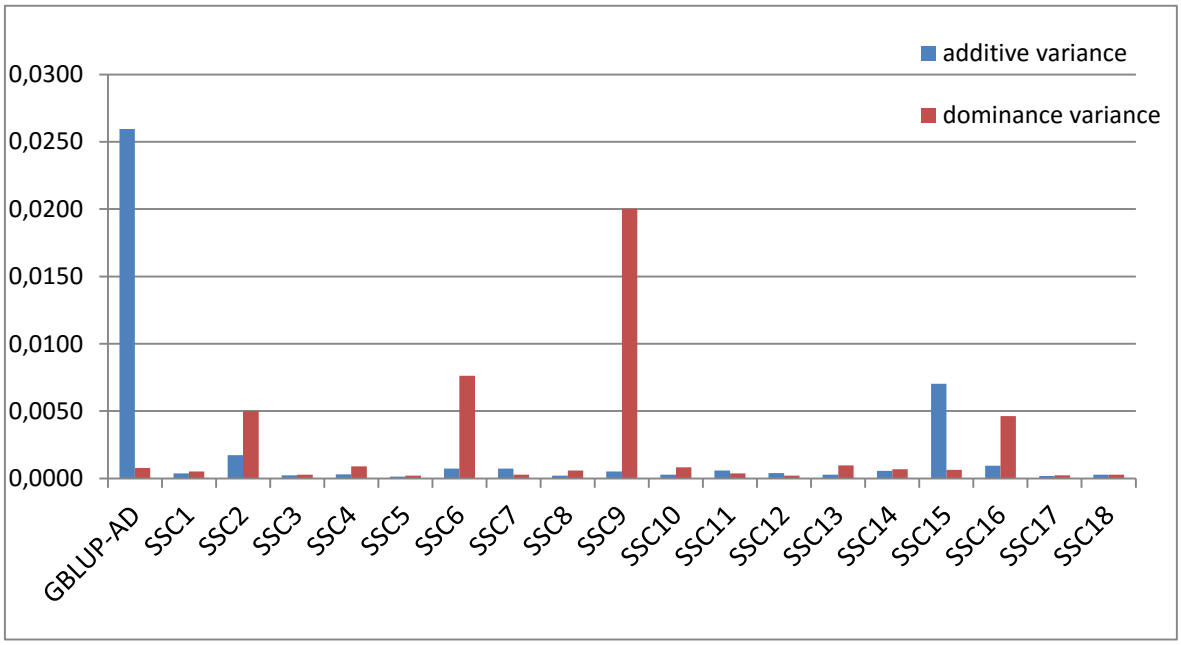


Figure 16. Genetic (σ_a^2 and σ_d^2) variance components for skatole estimated by GBLUP-AD for whole genome and regional heritability approach

Analysis of dominance for androstenone

Results in table 10, shows the estimates of additive, dominance, dam and residual variances using whole genomic relationship matrix for androstenone. Results show that the substantial amount of dominance variances was estimated for androstenone ($P<0.05$, from the LRT). On the whole genome level, estimated dominance variance with GBLUP-AD was 0.063 while its proportion in total phenotypic variance was 13%. Additive genetic variance estimated with GBLUP-A (0.14) was higher than with GBLUP-AD (0.11) as well as their proportions in the total variance, 28% and 22.5%, respectively. Narrow sense heritability of 0.28 was higher for GBLUP-A than a value of 0.23 for GBLUP-AD. However, estimate of broad sense heritability was higher in GBLUP-AD (0.36) model than the narrow sense heritability (0.23), as more genetic variance was explained by including dominance effects (0.06). Broad sense heritability estimated by GBLUP-D (0.35) model was close to the estimate of GBLUP-AD (0.36) what shows that the substantial variations due to dominance genetic effects present in this dataset. Dam genetic variance estimated with GBLUP-AD accounted for 7% of total phenotypic variance and was slightly lower compared to GBLUP-A (9%). For androstenone, likelihood ratio tests are shown in the table 10. These tests compare the log likelihoods of the two models and tests whether this difference is statistically significant or if the model with additional variance component provides more suitable model. In this particular dataset, adding dominance effects in the model have shown significant improvement as the likelihood ratio (LRT=3.44) exceeded threshold (2.71) for P value of 0.05. Additionally, when additive genomic relationship matrix was included, significant model improvement (LRT=17.65) was also achieved with exceeding the same critical value for 0.05.

In table 12, ratios of chromosomal additive (σ_a^2) and dominance (σ_d^2) genetic variances in total additive and dominance genetic variance, respectively, and significance of the LRT for androstenone are shown. Analyses of chromosomal variance components when fitting individual autosomal chromosome for androstenone showed that no significant additive genetic variation could be assigned to any one particular chromosome. The highest proportions of chromosomal additive genetic variance in total additive genetic variance were estimated on chromosome 15 (16%) and chromosome 5 (2%), but the values were not significant. Proportions of chromosomal dominance genetic variance in total dominance genetic variance were highest for chromosome 11 (27%), chromosome 10 (18%), chromosome 9 (11%) and chromosome 16 (9%), but only for the chromosome 11 the LRT was statistically significant. The estimates of fractions of dominance variance across chromosomes over the multiple analyses were reasonably consistent summing to 0.86. The estimates of maternal effect variances didn't change between the models and was similar across all autosomes (~0.04). These variances were not reported in tables.

Analysis of dominance for skatole

Results in table 11, shows the estimates of additive, dominance, dam and residual variances using whole genomic relationship matrix for skatole. Additive genetic variances estimated with GBLUP-A (0.026) and GBLUP-AD (0.026) on the whole genome level were almost the same, as well as estimated narrow sense heritabilities, 0.055 and 0.054, respectively. Due to the specific study design and selection of animals based on the skatole levels, heritability estimates were biased and consequently deflated. Dominance variances were very low in all models used. Dominance genetic variance accounted for 1.2% of the total phenotypic variation with GBLUP-D and only 0.01% with GBLUP-AD. Although the dominance variance was low, broad sense heritability estimated with GBLUP-D was 0.01, what indicates the presence of certain level of dominance deviations. An estimate of broad sense heritability was higher with GBLUP-AD (0.056), therefore certain improvement of the model is achieved by fitting dominance effects compared to GBLUP-A (0.054). Dam genetic effects estimated were very small numbers close to zero in all models used.

For skatole in this particular dataset, adding dominance effects in the model did not show significant improvement with the LRT=0.08. When additive genomic relationship matrix was included in the model, significant improvement was achieved as the LRT=3.6 surpassed the threshold of $p>0.05$.

Regional approach analyses of chromosomal variance components for skatole are shown in table 13, with the ratios of chromosomal additive (σ^2_g) and dominance (σ_d^2) genetic variances in total additive and dominance genetic variance, respectively, and significance of the LRT. Proportion of chromosomal additive genetic variation was highest on chromosome 14 (22%) and chromosome 15 (20%), but the values were not significant. However, much higher dominance genetic effects were captured on a chromosomal level. Significant model improvement was achieved by including dominance effects of chromosome 9, where the

proportion of chromosomal dominance genetic variance in total dominance variance was 96%. It is worth noting that only for SSC9 was there evidence of dominance variance and that in this case any remaining estimate was very small. The small total dominance variance component results in the fractions of variance explained by chromosomes and becoming very sensitive to the sampling errors and this is evident by the sum over the chromosomes explaining the estimated total variance many times over. Although not significant, highest chromosomal dominance variance estimated after SSC9 was with the chromosome 6 (88%), 2 (87%) and 16(87%). Other chromosomal dominance genetic variances were also not significant, but could indicate the presence of dominance genetic effects in this particular dataset.

6. DISCUSSION

Accuracy of GBLUP and Bayesian regression based methodologies

To our knowledge, this is the first study to test the different methodologies for genomic evaluation of androstenone and skatole concentrations in the fat of slaughtered male pigs, two traits that are the directly related to the occurrence of boar taint. It was shown that training data by using all markers simultaneously in genomic evaluations (Meuwissen et al., 2001) was giving better accuracies than using detected QTL. In the case of androstenone concentration the accuracies obtained from GBLUP or a series of Bayesian methods were very similar. In contrast, for skatole concentration, where it has been established that a large QTL is segregating within this population explaining 77% of the genetic variance (Rowe et al., 2014), Bayesian methods that provide for only a subset of SNP having large effects gave more accurate predictions than GBLUP. However such a benefit would not exist for breeds in which this QTL is not segregating.

The design, focusing primarily on selection within full-sib families, had the objective of increasing the power of identifying QTL in GWAS studies by boosting the value of long-term LD in obtaining marker-QTL associations and reducing the emphasis on LD arising from more recent family structure. Luan et al., (2012) showed that in some populations the more recent family structure can be captured by using linkage analysis to construct relationships, and that this can account for most of the achieved accuracy from genomic evaluation. The design has immediate consequences on the results presented as the selection will introduce biases into the estimates of predicting the phenotype and the estimates of heritability, whether genomic or otherwise (Daetwyler et al., 2008); and consequently for the estimates of accuracy for predicting breeding values as this prediction uses both of these

parameters. However, the ranking of the methods would not be expected to be affected by the selection on skatole concentration *per se*. Had sampling been at random from the population the expectation would be that accuracies (as reported from a cross-validation using such data) would reduce, since less-informative families would have been used. However the summary of the methods as stated at the start of the discussion would be expected to remain valid as the amount of population-wide data increased as the QTL-SNP associations would emerge more strongly.

As outlined earlier boar taint provides challenges for the breeder in that it is age- and sex-limited and destructive to measure directly. Initial attempts using selection on indirect traits, such as concentrations in the blood or size of the sex glands were less successful than anticipated. The genetic reasons for this relative failure came from initial heritability estimates that proved to be over-optimistic, and some unfavourable genetic correlations (Sellier et al., 2000; Sellier and Bonneau, 1988; Willeke et al., 1980). Reducing the expression of boar taint is expected to be associated with reduced androstenone concentrations in fat and blood, but since androstenone is synthesised together with other steroids, such as androgens and estrogens (Robic et al., 2011) selection against taint resulted in lengthening the time to sexual maturity in pigs with low androstenone levels. However, skatole appears in fat through a relatively short metabolic pathway (Zamaratskaia and Squires, 2009), which reduces the number of network interactions that may occur, and empirically a reduction in skatole has not been associated with a negative effect on sex hormones. Therefore skatole seems a more promising trait for utilizing in selection.

The results from this study advance the opportunities for selection against the expression of taint since it demonstrates that genomic predictions, simultaneously utilising all SNP, will offer opportunities to select against expression of taint that overcomes the age- and sex-

limitations and the destructiveness of measuring the trait. Furthermore these accuracies will increase as more data is obtained for training these genomic predictors. However the results do not address the remaining barrier to implementing genomic evaluations in practice, which are the uncertain and possibly unfavourable genetic correlations of the expression of boar taint with other traits of value.

Therefore, 3 plausible approaches could be suggested for practical application of genomics to reduce boar taint, although for all approaches described below more data collected population-wide should be obtained to validate findings and further improve accuracy. The first approach is to develop and utilise Bayesian models for skatole concentrations, which may be more free from unfavourable correlations than androstenone concentrations (Moe et al., 2009), and is considered to have the greater impact on customer acceptability (Bonneau and Squires 2004; Lee at al., 2005). This approach assumes the large QTL explaining substantial genetic variance is segregating in the population and the benefit arises from the possibility of obtaining greater accuracy from the genotyping that is conducted. Alongside this model, androstenone could be included as a trait in routine GBLUP evaluations to accumulate more information on genetic correlations. The second approach would be to use GBLUP for both androstenone and skatole, which allows more routine evaluations and easier to implement although losing accuracy in skatole concentration, which was estimated as 5% in the data presented. In both these approaches the accumulation of data would clarify the genetic correlations with other traits of value.

The third approach builds on either of the first two and would be to use the genomic predictors for skatole and androstenone concentrations, obtained from the accumulating field data, to be treated as a trait in more detailed studies carried out in elite populations. As the genomic predictor would be the selection criterion it is appropriate that it is the genetic

correlation with the predictor that is estimated rather than any other indirect measure of taint expression; moreover these genomic predictors are traits of accuracy 1 and hence correlations would be estimated more accurately within the less-numerous elite population. This is analogous to the use of BLUP EBV as an indication of potential correlated responses. Information on correlations of the 'marker-accessible' boar taint with other relevant traits would then allow incorporation of the genomic predictor into the selection index. Such an approach fulfils one of the long-term aspirations of genomics by utilising field records from lower down the pyramid to provide haplotypes for direct selection at the tip of the pyramid.

Effects of dominance on androstenone

The genetic architecture of quantitative traits is composed of three components and their interactions: additive (V_a), dominance (V_d) and epistatic (V_e) variance (Bulmer, 1985). Narrow sense heritability, defined as ratio of additive and total phenotypic variance, is considered as most important factor in transmission of trait from one generation to another. Previous genetic evaluations for boar taint related compounds, androstenone and skatole didn't account for genetic effects other than additive. As shown in some studies (Sellier et al., 2000), evaluations for boar taint related compounds had a low accuracy. To some extent, that loss of accuracy probably depends on non-additive genetic effects and animals with dominance relationships included in studies (Misztal, 1997). Results from this study verified the contribution of dominance variance in total phenotypic variation. For androstenone in this particular dataset, GBLUP-AD captured substantial dominance genetic variances as this selected population was comprised mainly of full sibs, where the expected degree of dominance relationships is high. The predictive ability of GBLUP-AD was improved as the estimates of heritability increased when dominance effects were included in the model. Estimates of broad sense heritability for androstenone were higher (GBLUP-AD H^2=0.36)

than the narrow sense heritability (GBLUP-A h^2=0.28). The dominance genetic variance accounted for 13%, while the additive accounted for 28% of the total phenotypic variation. LR test showed that fitting only dominance genomic relationship matrix enabled to explain more variance than the model with fitting only additive genomic relationship matrix. In the study of Da et al., (2014) with simulated data, prediction of total genetic values including both additive and dominance effects was more accurate than the prediction with only additive model. Similar results were achieved (Su et al., 2012) on the real dataset of Danish Duroc pigs where models with included dominance genetic effects had improved accuracy as well as slightly improved unbiasedness of prediction.

This research is based on pig dataset of purebred Danish Landrace population, therefore by selecting appropriate crossbreeding design an extra response could be achieved. Moghadar and van der Werf, (2014) in the study on purebred and crossbred sheep genotyped with 50K chip, showed significant model improvement by fitting dominance effects and higher accuracy of genomic breeding values in crossbreds. In pig production, crossbreds are the final product, therefore genotypic information from purebred parents could allow the selection of their crossbred offspring (Su et al., 2012).

Maternal or dam effects signify that dam had an influence on offspring performance. Those effects are completely environmental for the offspring, but they could have both genetic and environmental components (Falconer and Mackay, 1996). If the maternal additive genetic effects are present, they could affect the estimates in the model the direct additive genetic effects as they are transmitted together on the following generation. In this study, maternal or dam effect accounted for 7% of the total phenotypic variation for androstenone.

Effects of dominance on skatole

The results from the present study demonstrated that the trait architecture for androstenone and skatole is different. Androstenone is a steroid hormone for which studies suggested it to be under control of polygenic effects (Rowe et al., 2014), whereas skatole is a product from the intestinal microbial activity for which it is suggested to be major gene control. The majority of the studies that tried to estimate heritabilities, did not account for non-additive genetic effects in their models. In this research, selection was based on the skatole concentrations which affected the within family variation, therefore the estimates of heritabilities (0.05 with GBLUP-AD) were much lower than previous reports (Tajet et al., 2006). Including the dominance genetic effects in the model for skatole showed only slight improvement in the model fit. However, proportion of dominance variance in GBLUP-D model was 0.01 what suggests the certain level of dominance deviations.

Regional approach for estimation of heritabilities attributed to each chromosome, gave us further insight into the genetic basis of skatole. The research of Yang et al., (2010) on the human data demonstrated that the variance of single major gene can be captured by SNP-s from the whole genome and by regional approach using SNP-s from chromosomes separately and is independent of the effects sizes. This analysis showed higher proportion of additive genetic variance than dominance using whole genomic relationship matrix. However, regional approach showed higher proportion of dominance variance attributable to each chromosome. For chromosomes 9, 6, 2, and 16, ratios of dominance variance to total phenotypic variance was 4%, 1.3%, 1.1% and 1.1%, respectively. These results show that the genome partitioning methods could further help in clarifying the genetic architecture of the trait.

7. CONCLUSION

Main subject of this thesis is to provide new knowledge necessary for implementing genomic information in breeding programs against boar taint in commercial production.

Genomic evaluation methodologies are newly used techniques in animal breeding for the estimation of genetic merit of selection candidates. In order to choose optimum and most accurate methodology, genetic architecture of quantitative traits stands as an important factor.

For this dataset of a commercial Danish Landrace population, different ranges of accuracies are calculated using different methodologies of genomic evaluation against boar taint. For androstenone concentration, GBLUP and regression based methodologies perform with equal accuracy in predicting phenotypes, which was anticipated as prior evidence suggests genetic variance is not dominated by a few QTL. In contrast, when predicting skatole concentrations, Bayesian methodologies had greater accuracy than GBLUP, consistent with a large QTL known to be segregating in this population. The barriers to cost-effective genetic selection against boar taint, arising from the age- and sex-limitations and destructiveness of measuring boar taint can be removed using genomic evaluations, subject to developing a training set of adequate size. The development of predictors from field data can also assist removing uncertainties over unfavourable genetic correlations between boar taint and other traits of value, by utilising the genomic predictors in more detailed studies within elite populations.

In addition, including the non-additive genetic effects could further improve the accuracy of evaluations. Results for both androstenone and skatole using the whole genomic or regional

approach, suggest that dominance genetic effects affect their concentrations. Since the crossbreds are the end product in pig production, genomic predictors from purebreds and their crossbred offspring could help in selecting purebred candidates. Moreover, selection designs such as specific mate allocation could also provide an additional response. Based on these analyses, it is demonstrated that the methods of including dominance genomic relationship matrix provide additional source of genetic variation and at the same time feasible approach in genomic evaluations. The results obtained from this study demonstrate such solutions are worthwhile considering in national breeding strategies to address the need for reliance on castration.

As an additional remark of this research, certain points should be mentioned. Genomic predictors provide higher accuracies of GEBVs that surpasses drawbacks related with age, sex and late measuring. However, undesirable genetic relationship of androstenone with reproductive traits still stands as a barrier in selection against boar taint so more research should be conducted and focused on accurate dissection of these relationships. In that context, additional research should be performed in picking and accurate measuring of reproductive traits for boars and gilts that could negatively affect the selection. Recent studies already confirmed that selection against skatole should not affect some reproductive traits in boars, such as ejaculate quality and embryo survival (Strathe et al., 2014). However, as the previous selection (Willeke et al., 1987) negatively affected both boar and gilt reproduction, these parameters should be investigated in both sexes and other breeds so that the selection could be efficiently implemented.

8. REFERENCES

1. Aluwé M, Tuyttens F A M, Bekaert K M, Smet De S, Brabander D L De, Millet S (2012: Evaluation of various boar taint detection methods. Animal (2012), 6:11, pp 1868–1877
2. Aldal I., Andresen Ø., Egeli A. K., Haugen J. E., Grødum A., Fjetland O., Eikaas J. L. H.. 2005. Levels of androstenone and skatole and the occurrence of boar taint in fat from young boars. Livest. Prod. Sci. 95:121–129.
3. Amin N., van Duijn C. M., Aulchenko Y. S.. (2007) A genomic background based method for association analysis in related individuals. PLOS One 2: e1274.
4. Andersson, H.K., Andersson, K., Zamaratskaia, G., Babol, J., Rydhmer, L. and Lundström, K., 2003. The effect of single-sex or mixed sex raising of entire male and female pigs on production results, meat quality, skin damage and boar taint. EAAP Working Group on Production and Utilisation of Meat from Entire Male Pigs, Dublin, Ireland, 13-14 November 2003
5. Andresen, O. (1974) Development of a Radioimmunoassay for 5-andro-16-en-one in Pig peripheral Plasma. Acta Endocrinologica 76, 377-387
6. Arasta, P; Chen, Y; Zhang, Y; Kerr, R; Graser, H-U; Luxford, B and Moran, C (2007): "Cyp21 as a candidate gene for androstenone on boar taint." Proceedings of the 17th Conference of the Association for the Advancement of Animal Breeding and Genetics. Armidale, Australia. September 24-26. pp.65-68.
7. Babol J. and Squires EJ 1995. Quality of meat from entire male pigs. Food Research International 28, 201–212.
8. Babol J, Squires EJ and Lundstrøm K (1998): Relationship between oxidation and conjugation metabolism of skatole in pig liver and concentrations of skatole in fat. Journal of Animal Science 76, 829–838.

9. Babol J., Squires E. J. and Lundstrom K. (1999) Relationship between metabolism of androstenone and skatole in intact male pigs. Journal of Animal Science 77, 84-92.
10. Banoglu E, Jha GG, King RS: Hepatic microsomal metabolism of indole to indoxyl, a precursor of indoxyl sulfate. Eur J Drug Metab Pharmacokinet 2001, 26:235-240.
11. Bonneau, M., Dufour, R., Chouvet, C., Roulet, C., Meadus, W. and E.J. Squires (1994) The effects of immunization against luteinizing hormone-releasing hormone on performance, sexual development, and levels of boar taint-related compounds in intact male pigs. Journal of Animal Science 72, 14-20.
12. Bonneau, M., Desmoulin, B., and Dumont, B. L., 1979. Production de viandes de porc mâles entiers ou castrés: efficacité alimentaire et composition corporelle chez les races hypermusclées. Annales de Zootechnie, 28, 53-72
13. Bonneau M., Squires J., (2004): Boar taint: Causes and measurement. In Encyclopedia of Meat Sciences. Eds W.K. Jensen, C. Devine & M. Dikemann, Elsevier, Oxford, ISBN 0-12-464970-X, pp.91-96 (I-S)
14. Bonneau M, LeDenmat M, Vaudelet JC, VelosoNunes JR, Mortensen AB and Mortensen HP (1992): Contributions of fat androstenone and skatole to boar taint: I. Sensory attributes of fat and pork meat. Livestock Production Science 32, 63–80.
15. Bonneau M, Carrie´ -Lemoine J, Prunier A, Garnier DH and Terqui M 1987. Age-related changes in plasma LH and testosterone concentration profiles and fat androstenone content in the young boar. Animal Reproduction Science 15, 241–258.
16. Bonneau, M., Kempster, A.J., Claus, R., Claudi-Magnussen, C. Diestre, A., Tornberg, E., Walstra, P., Chevillon, P. Weiler, U., and Cook, G. L., 2000. An international study on the importance of androstenone and skatole for boar taint: I.

Presentation of the programme and measurement of boar taint compounds with different analytical procedures. Meat Science, 54, 251-259.

17. Brennan, J.J., Shand, P.J., Fenton, M., Nicholls, L.L., and Aherne, F.X., 1986. Androstenone, androstenol and odour intensity in backfat of 100- and 130-kg boars and gilts. Canadian Journal of American Science, 66, 615-624
18. Brooks R.I., Pearson A.M., (1986): Steroid hormone pathways in the pig, with special emphasis on boar odor: a review, J. Anim. Sci. 62 632–645.
19. Burkett, Jeremy L., "The effect of selection for intramuscular fat on fatty acid composition in Duroc pigs" (2009). Graduate Theses and Dissertations. Paper 10539. Iowa State University
20. Chen W, Forrest JC, Peng IC, Pratt DE, Judge MD. Palatability of prerigor cooked boar meat. J Anim Sci. (1993): 71:645–650.
21. Claus R, Weiler U and Herzog A (1994): Physiological aspects of androstenone and skatole formation in the boars – a review with experimental data. Meat Science 38, 289–305.
22. Daetwyler H.D., Villanueva B., Woolliams J.A. (2008): Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. PLoS ONE 3(10): e3395.
23. Daetwyler, H.D., Pong-Wong R., Villanueva B. and Woolliams J. A. (2010) The impact of genetic architecture on genome-wide evaluation methods. Genetics 185, 1021-1031.
24. Davis SM and Squires EJ 1999. Association of cytochrome b5 with 16-androstene steroid synthesis in the testis and accumulation in the fat of male pigs. Journal of Animal Science 77, 1230–1235.

25. Davoli, R. i Braglia, S. (2008): Molecular approaches in pig breeding to improve meat quality, Briefings In Functional Genomics And Proteomics. Vol 6. No 4. 313-321.
26. de Roest K., Montanari C., Fowler T. and Baltussen W. (2009) Resource efficiency and economic implications of alternatives to surgical castration without anaesthesia. Animal 3:11, 1522–1531
27. de Lange, C.F.M. i J. Squires (1995): Entire males vs. castrates – financial benefits to the producer. Ontario Swine Research Review. 41-44.
28. Diaz GJ and Squires EJ 2000. Metabolism of 3-methylindole by porcine liver microsomes: responsible cytochrome P450 enzymes. Toxicological Sciences 55, 284–292.
29. Doran, E., Whittington, F.W., Wood, J.D., McGivan, J.D. (2002): Cytochrome P450IIE1 (CYP2E1) is induced by skatole and this induction is blocked by androstenone in isolated pig hepatocytes, Chemico-Biological Interactions 140, 81–92.
30. Doran E, Whittington FM, Wood JD and McGivan JD 2004. Characterization of androstenone metabolism in pig liver microsomes. Chemico-Biological Interactions 147, 141–149.
31. Duchemin S. I., Colombani C., Legarra A., Baloche G., Larroque H., Astruc J.M., Barillet F., Granié C. Robert, and Manfredi E. (2012): Genomic selection in the French Lacaune dairy sheep breed. J. Dairy Sci. 95 :2723–2733
32. Duijvesteijn, N., E. F. Knol, J. W. Merks, R. P. Crooijmans, M. A. Groenen, H. Bovenhuis, and B. Harlizius. 2010. A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. BMC Genet. 11:42.

33. Dunshea FR, Colantoni C, Howard K, McCauley I, Jackson P, Long KA, Lopaticki S, Nugent EA, Simons JA, Walker J and Hennessy DP (2001): Vaccination of boars with a GnRH vaccine (Improvac) eliminates boar taint and increases growth performance. Journal of Animal Science 79, 2524–2535.
34. EFSA (European Food Safety Authority) (2004): Welfare aspects of the castration of piglets. The EFSA Journal (2004) 91, 1-18
35. Engelsma, K., R. Bergsma, B. Harlizius, D. Ducro-Steverink, and E. Knol. 2007. Genetic parameters for components of boar taint and their relation with carcass quality and female fertility. EAAP 2007, Dublin, Ireland, August 26-29.
36. European Declaration on alternatives to surgical castration of pigs, 2011. http://ec.europa.eu/food/animal/welfare/farm/docs/castration_pigs_declaration_en.pdf
37. Falconer, D. S. and Mackay, T. F. C. (1996). Introduction to Quantitative Genetics. Longman, Essex, U.K., 4th edition.
38. FAO Biannual report on global food markets, 2013. http://www.fao.org/docrep/019/i3473e/i3473e.pdf
39. Font-i-Furnols M. (2012): Consumer studies on sensory acceptability of boar taint: A review Meat Science 92 (2012) 319–329
40. Fowler VR, McWilliam T and Aitken R 1981. Voluntary feed intake of boars, castrates and gilts given diets of different nutrient density. Animal Production 32, 357
41. Fredriksen, B., Font i Furnols, M., Lundström, K., Migdal, W., Prunier, A., Tuyttens, F. A. M. and Bonneau, M (2009): Practice on castration of piglets in Europe, Animal, 3: 1480-1487

42. Gilbert, A. N. and Wysocki, C. J., 1987. The smell survey. Results. National Geographics 172, 514-525

43. Gilmour A.R., Cullis B.R., Welham S.J., Thompson R. (2000): ASReml Reference manual. NSW Department of Agriculture, Orange. NSW. Agriculture Biometric 2000, 3:210.

44. Gispert, M., Oliver, M. A., Velarde, A., Suarez, P., Pérez, J and Font i Furnols, M. (2010) Carcass and meat quality characteristics of immunocastrated male, surgically castrated male, entire male and female pigs, Meat Science, 85: 664-670

45. Gregersen V.R., Conley L.N., Sorensen K.K., Guldbrandtsen B., Velander I.H., Bendixen C. (2012) Genome-wide association scan and phased haplotype construction for quantitative trait loci affecting boar taint in three pig breeds. BMC Genomics 2012, 13:22.

46. Greger D., Genetic marker for meat quality, growth, carcass and reproductive traits in livestock, patent WO 0/69882,Washington, DC (2000).

47. Grindflek E., Lien S., Hamland H., Hansen M.H., Kent M., van Son M. and Meuwissen Theo H.E. (2011) Large scale genome-wide association and LDLA mapping study identifies QTLs for boar taint and related sex steroids. BMC Genomics 2011, 12: 362.

48. Hansen LL, Mejer H, Thamsborg SM, Byrne DV, Roepstorff A, Karlsson AH, Hansen-Møller J, Jensen MT and Tuomola M 2006. Influence of chicory roots (Cichorium Intybus L.) on boar taint in entire male and female pigs. Animal Science 82, 359–368.

49. Haugen JE, Brunius C, Zamaratskaia G. (2012): Review of analytical methods to measure boar taint compounds in porcine adipose tissue: the need for harmonized methods. Meat Sci 90, 9-19.

50. Hayes, B.J.; Bowman, P.J.; Chamberlain, A.J., Goddard, M. E. (2009): Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science Vol. 92 No. 2
51. Hortos, M., Rius, M. A., De Vries, A., Lacoste, A., Gispert, M. and Diestre, A., 2000. In Proceedings 46th International Congress of Food Science and Technology, (Vol.1) (p. 98- 99), 27 August - 1 September 2000, Buenos Aires, Argentina
52. Ibañez-Escriche N. and O. Gonzalez-Recio, 2010: Promises, pitfalls and challenges of genomic selection in breeding programs. Spanish Journal of Agricultural Research 2011 9(2)
53. Kemper, K.E., Visscher, M.P., Goddard, M. E. (2012): Genetic architecture of body size in mammals. Genome Biology 2012, 13:244.
54. Le Mignon, G., Ianuccelli, N., Robic, A., Billon, Y., Bidanel, J.P., Larzul, C. (2010): Fine mapping of quantitative trait loci for androstenone and skatole levels in pig, Proceedings of 9th World Congress on Genetics Applied to Livestock Production (WCGALP).
55. Lee G.J., Archibald A.L., Law A.S., Lloyd S.,Wood J., Haley C.S. (2005) Detection of quantitative trait loci for androstenone, skatole and boar taint in a cross between Large White and Meishan pigs. Animal Genetics 36, 14–22.
56. Lee S. H., Goddard M. E., Visscher P.M., van der Werf J. H.J. (2010): Using the realized relationship matrix to disentangle confounding factors for the estimation of genetic variance components of complex traits Genetics Selection Evolution 2010, 42:22
57. Lin Z., LouY., Squires E.J., (2004): Molecular cloning and functional analysis of porcine SULT1A1 gene and its variant: a single mutation SULT1A1 causes a significant decrease in sulfation activity, Mamm. Genome 15, 218–226.

58. Lin, Z., Lou, Y., Squires, J.E. (2006): Functional polymorphism in porcine CYP2E1 gene: Its association with skatole levels, Journal of Steroid Biochemistry & Molecular Biology 99 231–237.

59. Luan T., Woolliams J.A., Odegard J., Dolezal M., Roman-Ponce S. I., Bagnato A. and Meuwissen T.H.E. (2012) The importance of identity-by-state information for the accuracy of genomic selection. Genetics Selection Evolution 2012 44:28.

60. Meat and Livestock Commission, 1989. Stotfold pig development unit, results from the first trial. Meat and Livestock Commission, Milton Keynes, England.

61. Meadus, W.J., Mason, J.I., and Squires E.J., 1993. Cytochrome P450c17 from porcine and bovine adrenals catalyses the formation of 5,16-androstadien-3α-ol from pregnenolone in the presence of cytochrome b5. The Journal of Steroid Biochemistry and Molecular Biology, 46:565-572

62. Meuwissen, T. H. E., Hayes, B. J. and Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819-1829

63. Meuwissen T. (2003): Genomic selection: The future of marker assisted selection and animal breeding. In Marker assisted selection: A fast track to increase genetic gain in plant and animal breeding? Pp. 54-59. Turin, Italy, 17-18 Oct. 2003. FAO, Rome.

64. Moe M., Lien S., Aasmundstad T., Meuwissen T.H.E., Hansen M.H.S., Bendixen C. and Grindflek E. (2009) Association between SNPs within candidate genes and compounds related to boar taint and reproduction. BMC Genomics 10:32

65. Moghaddar N. and van der Werf J.H.J (2014): Genomic estimation of additive and dominance genetic variance and their effect on the accuracy of genomic prediction of sheep. Proceedings, 10th World Congress of Genetics Applied to Livestock Production. 17 – 22 August 2014, Vancouver, Canada.

66. Møller J. H. and Andersen J. R. (1994) Boar taint – analytical alternatives. Fleischwirtsch 74 (9), 963-966

67. Moe M, Meuwissen T, Lien S, Bendixen C, Wang X, Conley LN, Berget I, Tajet H, Grindflek E (2007): Gene expression profiles in testis of pigs with extreme high and low levels of androstenone. BMC Genomics 8, 405.

68. Moe, M., Lien, S., Aasmundstad, T., Meuwissen, T.H.E., Hansen, M.H.S., Bendixen, C. and Grindflek, E. (2009). Association between SNPs within candidate genes and compounds related to boar taint and reproduction. BMC Genomics 10:32

69. Mortensen, A.B., and Sorensen, S.E., 1984. Relationship between boar taint and skatole determined with a new analysis method. 30th European Meeting of Meat Research Workers, Bristol, U.K.

70. Mortensen AB, Bejerholm C, Pedersen JK. Consumer test of meat from entire males, in relation to skatole in backfat. Proc 32nd European Meeting of Meat Research Workers, Ghent. 1986;p23–26.

71. Mrode, R. A., 2005 Linear Models for the Prediction of Animal Breeding Values, Ed. 2. CABI Publishing, Oxfordshire, UK.

72. Nelson D.R, Koymans L, Kamataki T, Stegeman J.J, Feyereisen R, Waxman D.J, Waterman M.R, Gotoh O, Coon M.J, Estabrook R.W, Gunsalus I.C, Nebert D.W. (1996): P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics. 1996, 6(1):1-42.

73. Nirea K.G., Sonesson A.K., Woolliams J. A. and Meuwissen T.H.E. (2012): Strategies for implementing genomic selection in family-based aquaculture breeding schemes: double haploid sib test populations. Genetics Selection Evolution 2012 44:30.

74. Payne AH, Hales DB: Overview of steroidogenic enzymes in the pathway from cholesterol to active steroid hormones. Endocrine Reviews 2004, 25:947-970.
75. Pearson AM, Ngoddy S, Price JF, Larzelere HE. Panel acceptability of products containing boar meat. J Anim Sci. 1971;33:26–29
76. Pedersen JK, Mortensen AB, Madsen A, Mortensen HP, Hyldegaard-Jensen J. [The influence of feed on boar taint in meat from pigs.] National Institute of Animal Science (NIAS), Denmark, Communication, 1986; No. 638: p4.
77. Pravilnik o minimalnim uvjetima za zaštitu svinja (NN119/10).
78. Quintanilla R, Demeure O, Bidanel JP, Milan D, Iannuccelli N, Amigues Y, Gruand J, Renard C, Chevalet C and Bonneau M 2003. Detection of quantitative trait loci for fat androstenone levels in pigs. Journal of Animal Science 81, 385–394.
79. Ramos A.M., Crooijmans R.P., Affara N.A., Amaral A.J., Archibald A.L., Beever J.E., Bendixen C., Churcher C., Clark R., Dehais P., Hansen M.S., Hedegaard J., Hu Z.L., Kerstens H.H., Law A.S., Megens H.J., Milan D., Nonneman D.J., Rohrer G.A., Rothschild M.F., Smith T.P., Schnabel R.D., Van Tassell C.P., Taylor J.F., Wiedmann R.T., Schook L.B., Groenen M.A. (2009) Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS One 2009, 4: e6524.
80. Rideout TC, Fan MZ, Cant JP, Wagner-Riddle C and Stonehouse P 2004. Excretion of major odor-causing and acidifying compounds in response to dietary supplementation of chicory inulin in growing pigs. Journal of Animal Science 82, 1678–1684.
81. Ritz Josip, (1996): Hrvatsko-engleski, englesko-hrvatski agronomski rječnik, Školska knjiga, Zagreb.

82. Robic A, Le Mignon G, Feve K, Larzul C, Riquet J (2011): New investigations around CYP11A1 and its possible involvement in an androstenone QTL characterised in Large White pigs. Genet Sel Evol 43, 15.
83. Robic, A., Larzul, C. i Bonneau, M. (2008): Genetic and metabolic aspects of androstenone and skatole deposition in pig adipose tissue: A review, Genet. Sel. Evol., 40:129-143.
84. Roca J, Parrilla I, Rodriguez-Martinez H, Gil MA, Cuello C, Vazquez JM and Martinez EA (2011): Approaches Towards Efficient Use of Boar Semen in the Pig Industry. Reprod Dom Anim 46 (Suppl 2), 79–83
85. Rowe S. J, Karacaören B., de Koning D.-J., Lukic B., Hastings-Clark N., Velander I., Haley C. S and Archibald A. L (2014) Analysis of the genetics of boar taint reveals both single SNPs and regional effects. BMC Genomics 2014 15:424.
86. Sather AP, Squires EJ, Jeremiah LE, Jones SDM, Schaefer AL. (1992): Meat quality and consumer acceptance of pork from entire males. Can J Anim Sci. 72:1014–1015.
87. Sellier P., Bonneau M., (1988): Genetic relationships between fat androstenone level in males and development of male and female genital tract in pigs, Journal of Animal Breeding and Genetics 105:11–20
88. Sellier P., Le Roy P., Fouilloux M.N., Gruand J., Bonneau M., (2000) Responses to restricted index selection and genetic parameters for fat androstenone level and sexual maturity status of young boars. Livestock Production Science 63 265–274.
89. Sinclair PA and Squires EJ 2005. Testicular sulfoconjugation of the 16-androstene steroids by hydroxysteroid sulfotransferase: its effect on the concentrations of 5alpha-androstenone in plasma and fat of the mature domestic boar. Journal of Animal Science 83, 358–365.

90. Sinclair PA, Gilmore WJ, Lin Z, Lou Y and Squires EJ 2006. Molecular cloning and regulation of porcine SULT2A1: relationship between SULT2A1 expression and sulfoconjugation of androstenone. Journal of Molecular Endocrinology 36, 301–311.
91. Skinner, T.M., Doran,E., McGivan, J.D., Haley, C.S. i Archibald, A.L. (2005): Cloning and mapping of the porcine cytochrome-p450 2E1 gene and its association with skatole levels in the domestic pig, Animal Genetics, 36, 417–422.
92. Skinner TM, Anderson JA, Haley CS and Archibald AL 2006. Assessment of SULT2A1, CYP2A6 and CYPC18 as candidate genes for elevated backfat skatole levels in commercial and experimental pig populations. Animal Genetics 37, 521–522.
93. R.J. Spelman, J. Arias, M.D. Keehan, V. Obolonkin, A.M. Winkelman, D.L. Johnson and B.L. Harris (2010): Application Of Genomic Selection In The New Zealand Dairy Cattle Industry. The 9th World Congress on Genetics Applied to Livestock Production (WCGALP) Leipzig, Germany. 1-6 Aug 2010.
94. Squires E.J. and Lundstrom K. (1997) Relationship between cytochrome P450IIE1 in liver and levels of skatole and its metabolites in intact male pigs. Journal of Animal Science 75, 2506–2511.
95. Squires E.J. and Bonneau M. (2004): Boar taint: Control. In Encyclopedia of Meat Sciences. Eds W.K. Jensen, C. Devine & M. Dikemann, Elsevier, Oxford, ISBN 0-12-464970-X, pp.97-103 (I-S)
96. Squires, E.J. i Schenkel, F.S. (2010): Managing Boar Taint: Focus on Genetic Markers, London Swine Conference Proceedings 2010, 99-102.
97. Tajet H, Andresen O and Meuwissen THE 2006. Estimation of genetic parameters for boar taint: skatole and androstenone and their correlations with sexual maturity. Acta Veterinaria Scandinavica 48(suppl. 1): S9.

98. Tambyrajah WS, Doran E, Wood JD and McGivan JD 2004. The pig CYP2E1 promoter is activated by COUP-TF1 and HNF-1 and is inhibited byandrostenone. Archives of Biochemistry and Biophysics 431, 252–260.

99. Terner MA, Gilmore WJ, Lou Y and Squires EJ 2006. The role of CYP2A and CYP2E1 in the metabolism of 3-methylindole in primary cultured porcine hepatocytes. Drug Metabolism and Disposition 34, 848–854.

100. Tuomola M., Riikka H., Pia K., Heikki M. and Lovgren T. (1997) Time-Resolved Fluoroimmunoassay for the Measurement of Androstenone in Porcine Serum and Fat Samples. Journal of Agricultural and Food Chemistry 45, 3529-3534

101. Turkstra, J.A., Zeng, X.Y., Diepen, J.T.M., van, Jongbloed, A.W., Oonk, H.B., Wiel, D.F.M., van and Meloen, R.H. (2002) Performance of male pigs immunised against GnRH is related to the time of onset of biological response. Journal of Animal Science 80, 2953-2959

102. Van Raden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F., Schenkel, F.S. (2009) Invited Review: Reliability of Genomic Predictions for North American Holstein Bulls. J. Dairy Sci. 92, 16-24

103. Visscher P.M. (2006): A Note on the Asymptotic Distribution of Likelihood Ratio Tests to Test Variance Components. Twin Research and Human Genetics Volume 9 Number 4 pp. 490–495

104. Wang, Chunkao; Habier, David; Wolc, Anna; Garrick, Dorian J.; Fernando, Rohan L.; Lamont, Susan J.; Dekkers, Jack C.M.; Kranis, Andreas; and Watson, Kellie A. (2013) "Application of Genomic Selection Using an Evenly Spaced Low-density Marker Panel in Broiler Chickens," Animal Industry Report: AS 659, ASL R2807.

105. Wellmann R., Preuß S., Tholen E., Heinkel J., Wimmers K. and Bennewitz J. (2013): Genomic selection using low density marker panels with application to a sire line in pigs. Genetics Selection Evolution 2013, 45:28

106. Wesoly R, Weiler U (2012): Nutritional Influences on skatole formation and skatole metabolism in the pig. Animals 2, 221-242.

107. Willeke H, Claus R, Pirchner F, Alsing W. A selection experiment against 5a-androst- 16-ene-3-one, the boar taint steroid, in adipose tissue of boars. Z Tierzuct ZuchtBiol. 1980;97:86–94.

108. Willeke, H., Claus, R., Müller, E., Pirchner, F., and Karg, H., 1987. Selection for high and low level of 5α-androst-16-en-3-one in boars. I. direct and correlated response of endocrinological traits. Journal of Animal Breeding and Genetics, 104: 64-73

109. Willeke H, Pirchner F. (1989): Selection for high and low level of 5-androst-16-en-3-one in boars. II. Correlations between growth traits and 5a-androstenone. J Anim Breed Genet.106: 312-–317.

110. Wood, J.D., and Enser, M., 1982. Comparison of boars and castrates for bacon production. 2. Composition of muscle and subcutaneous fat, and changes in side weight during curing. Animal Production, 35, 65-74

111. Wysocki, C.J., and Beauchamp, G.K., 1984. Ability to smell androstenone is genitacally determined. Proceedings of the Natural Academy of Science of the United States of America, 86, 7976-7978

112. Xue JL, Dial GD. Raising intact male pigs for meat (1997): Detecting and preventing boar taint. Swine Health and Production. 5(4):151–158.

113. Yost G.S., (1989): Mechanisms of 3-methylindole pneumotoxicity, Chem. Res. Toxicol. 2 273–279.

114. "Zakon o zaštiti životinja (NN135/06, 37/13, 125/13)"

115. Zamaratskaia G, Chen G and Lundstrom K (2006): Effects of sex, weight, diet and hCG administration on levels of skatole and indole in the liver and hepatic activities of cytochromes P4502E1 and P4502A6 in pigs. Meat Science 72, 331–338.

116. Zamaratskaia, G. and Squires, E.J. (2009): Biochemical, nutritional and genetic effects on boar taint in entire male pigs. Animal 3:1508-1521.

117. Zammerini D., Wood J.D., Whittington F.M., Nute G.R., Hughes S.I., Hazzledine M., Matthews K. (2012): Effect of dietary chicory on boar taint. Meat Science 91 (2012) 396–401

9. SUMMARY

Genetic selection against boar taint, which is caused by high skatole and androstenone concentrations in fat, is a more acceptable alternative than the current practice of castration. Genomic predictors offer an opportunity to overcome the limitations for such selection caused by the phenotype being expressed only in males at slaughter, and this study evaluated different approaches to obtain such predictors.

Samples from 941 pigs were included in a design which was dominated by 421 sib pairs, each pair having an animal with a high and a low skatole concentration below this threshold (≥0.3 μg/g). All samples were measured for both skatole and androstenone and genotyped using the Illumina SNP60 porcine beadchip (Illumina, San Diego, CA) for 62,163 Single Nucleotide Polymorphisms (SNP). The accuracy of predicting phenotypes was assessed by cross-validation using six different genomic evaluation methods, GBLUP and five Bayesian methods. In addition, this was compared to those obtained from predictions using only QTL that showed genome wide significance.

The range of accuracies obtained by different prediction methods was narrow for androstenone, between 0.29 (Bayes Lasso) and 0.31 (Bayes B), and wide for skatole, between 0.21 (GBLUP) and 0.26 (Bayes SSVS). Relative accuracies corrected for h2, were 0.54-0.56 and 0.75-0.94 for androstenone and skatole, respectively. The whole genome evaluation methods gave greater accuracy than using QTL alone (one SNP for androstenone and one SNP for skatole).

Also, the dominance genetic variation was ignored in national evaluations, so we assessed the dominance genetic variance for androstenone and skatole. This was achieved by constructing the dominance genomic relationship matrix from SNP information.

For androstenone in this dataset, GBLUP with dominance effects included captured substantial ratios of the dominance genetic variances (13%) in total variation. For skatole,

more dominance genetic variance was captured by regional chromosomal heritability approach, particularly on chromosome 9, where the proportion of chromosomal dominance genetic variance in total dominance variance was 96%. The results obtained demonstrate that GBLUP for androstenone is the simplest genomic technology to implement and was also close to the most accurate method. More specialised models may be preferable for skatole while the dominance genomic relationship matrix provide additional source of genetic variation for both traits, therefore it is worthwhile to consider in genomic evaluations.

10. SAŽETAK

Nerastovsko svojstvo je pojava neugodnog mirisa i okusa mesa koja nastaje radi prekomjernog nakupljanja skatola i androstenona u masnom tkivu kod nekih muških svinja. Genetska selekcija je dugoročno promatrano prihvatljivije rješenje od kastracije koja se trenutno provodi kao preventivna mjera protiv suzbijanja nerastovskog svojstva. Genomski markeri odnosno prediktori, pružaju nove mogućnosti u prevladavanju dosadašnjih ograničenja u selekciji svinja protiv spomenutog svojstva stoga je cilj ovog istraživanja istražiti mogućnosti njihove primjene.

U istraživanje je uključen 941 nerast, od kojih je 421 uparenih srodnika po ocu i majci, odnosno braće, te je odabiranjem cilj bio postići da svaki nerast s visokom razinom skatola ima srodnika po ocu i majci iz istog legla s niskom razinom skatola (≥0.3 ug / g). Svim uzorcima su osim razine skatola utvrđene i razine androstenona. Životinje su genotipizirane prema 62163 SNP-a (engl. Single Nucleotide Polymorphisms) koristeći Illumina SNP60k porcine beadchip (Illumina, San Diego, CA). Analiza unakrsne provjere je izvršena s ciljem uspoređivanja točnosti procjene GBLUP metode i pet Bayes metoda na osnovi regresije u procjenjivanju nepromatranih fenotipskih svojstava. Pored toga, izvršene su dodatne procjene koristeći samo jedan QTL koji je pokazao značajan utjecaj na razini cijelog genoma.

Raspon ostvarenih točnosti koristeći različite metode procjena je bio uzak za androstenon, između 0,29 (Bayes Lasso) i 0,31 (Bayes B), te širi za skatol, između 0,21 (GBLUP) i 0,26 (Bayes SSVS). Relativne točnosti korigirane s prosječnim H2, iznosile su 0,54 - 0,56 za androstenon i 0,75-0,94 za skatol. Metode procjene na razini cijelog genoma postigle su višu točnost od metode sa samo jednim statistički značajnim QTL-om (jedan statistički značajan SNP za androstenon i jedan za skatol).

S obzirom na to da se učinci dominantnosti u ukupnoj genetskoj varijabilnosti uglavnom nisu promatrali u nacionalnim uzgojnim programima evaluacije, dodatni je pristup testiran za procjenu dominantne genetske varijance za androstenon i skatol.

Za androstenon je na ovom skupu podataka, GBLUP s uključenim dominantnim učincima u modelu procijenio značajne omjere dominantnih genetskih varijanci (13%) u ukupnoj varijabilnosti. U slučaju skatola, više dominantnih genetskih varijanci je procijenjeno pristupom regionalnog heritabiliteta, i to najviše na kromosomu 9, gdje je udio kromosomske regionalne dominantne varijance u ukupnoj dominantnoj varijanci iznosio 96%. Dobiveni rezultati pokazuju da je GBLUP najjednostavnija metoda genomske procjene za androstenon, ujedno je lako provediva te jedna od najtočnijih metoda. U slučaju skatola, potrebno je kreirati prilagođene modele koji mogu postići značajno bolje rezultate. Također, metode koje uključuju dominantne genomske matrice srodstva pružaju dodatne izvore genetske varijabilnosti i istodobno nude novi i primjenjivi pristup u genomskim procjenama u okviru nacionalnih uzgojnih programa.

11.APPENDIX

CURRICULUM VITAE

Boris Lukić was born on August 5, 1983 in Osijek, Croatia. After completion of elementary and grammar school, in 2002 he was enrolled at the Faculty of Agriculture in Osijek, graduate studies of Animal husbandry. During his studies he was awarded the scholarship of the Ministry of Science for being among the 10% of the most successful students. He graduated in 2009 with the thesis: "Investigation of meat quality of hybrid pigs" and became a graduate engineer of agriculture, program zootechnics. While studying he spent several months doing the research internships in Germany at the National Agricultural Research Center "FAL Braunschweig" and in Poland at AgroTak Group cattle farm. Additionally, he worked as the local coordinator for the world's largest association for student exchange – I.A.E.S.T.E. Several times has Boris been a participant of the Festival of Science with lectures and workshops, and he has participated in Summer school at the Mediterranean Institute for Life Sciences "MedILS" Split, Croatia. In 2006 he won the Rector's Award as the most successful student. In 2009 he started his scientific career as a research assistant at the Faculty of Agriculture, department of Special Zootechnics and that same year he enrolled at the aforementioned faculty in postgraduate study "Agricultural Sciences" majoring in "Animal breeding". As a PhD student, he has attended advanced courses in statistics. Within the Erasmus Lifelong Learning program he visited the Roslin Institute of Edinburgh University, where he later spent 6 months doing a PhD research on a project about the prediction of genomic breeding values and genomic selection against boar taint. He actively participates as a research associate on the project of the Ministry of Science, Education and Sports "Early prediction of pig carcass and meat quality“ and VIP project of the Ministry of Agriculture, Fisheries and Rural Development, "The quality of pig carcasses as a function to increase the competitiveness of pig production." As a young researcher at the Faculty, he teaches students practical courses and seminars on modules of undergraduate and graduate programs, such as “Animal products”, “Animal products – Quality insurance” and “Genomic in zootechnics”. His research

focuses on how to improve the quality of animal products using different genomic approaches, such as candidate gene or genome wide approaches.