DANIELE BOTELHO DINIZ MARQUES

GENETIC PARAMETERS AND GENOMIC ANALYSIS OF SEMEN QUALITY AND FERTILITY TRAITS IN PIGS

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2017

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

M357g
2017

Marques, Daniele Botelho Diniz, 1988-
Genetic parameters and genomic analysis of semen quality and fertility traits in pigs / Daniele Botelho Diniz Marques. - Viçosa, MG, 2017.
viii, 107f. : il. (algumas color.) ; 29 cm.

Orientador: Paulo Sávio Lopes.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Suínos - Melhoramento genético. 2. Suínos - Inseminação artificial. 3. Marcadores genéticos. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-graduação em Zootecnia. II. Título.

CDD 22 ed. 636.40821

DANIELE BOTELHO DINIZ MARQUES

GENETIC PARAMETERS AND GENOMIC ANALYSIS OF SEMEN QUALITY AND FERTILITY TRAITS IN PIGS

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: September 4th, 2017.

John Wilhelmus M. Bastiaansen

Maria Leonarda Wilhelmina Johanna Broekhuijse

Fabyano Fonseca e Silva
(Co-advisor)

Simone Eliza Facioni Guimarães
(Co-advisor)

Paulo Sávio Lopes
(Advisor)

ACKNOWLEDGMENTS

I first would like to thank God for all protection and strength that allowed me to conclude my studies.

My parents Vera Lúcia and José Carlos for all love and encouragement, for being my support every moment I needed.

My husband Renato for everything, especially his support, companionship, patience, love and help since my under graduation and for being my company during my last period in the Netherlands.

All my family for always cheer with my achievements.

My mother-in-law Cor Maria, my father-in-law José Renato, my brothers-in-law Adriano and Mário and their wives Rosana and Fernanda for all affection, assistance and good moments we had together.

My friends, in special my bridesmaids Ana Paula, Stéphane and Juliana for the affection and friendship.

Universidade Federal de Viçosa and Animal Science Department for support since my under graduation.

Breeding and Genetics and Animal Science Graduate Programs of Universidade Federal de Viçosa for giving me support to conclude my doctorate degree and for the lessons received from the professors and courses offered.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support with scholarships.

My advisor since under graduation, Professor Paulo Sávio Lopes, for all attention, shared knowledge and incentive.

My co-advisor Professor Fabyano Fonseca e Silva for patience, pleasure in helping, advices and help with the analyses.

My co-advisor Professor Simone Eliza Facioni Guimarães for the help, incentive and confidence in my work.

Professor Renata for the help and friendship since my under graduation.

Colleagues from UFV Animal Breeding Group, for the support and shared knowledge.

Friends from LABTEC: Carol, Darlene, Ivan, Aline, Karine, Débora, Laene, Lucas, Margareth, Mariana, Márcio, Natalia, Pamela, Diego, Haniel, Susana and Walmir for all companionship, good moments and support during these years.

Wageningen University & Research, for allowing me do part of my doctorate degree at the University, providing all support that I needed.

WUR Animal Breeding and Genetic Group for lessons, good discussions and support during my training period.

John Bastiaansen, my advisor during my training period at WUR, for all help, lessons, good discussions, tips, patience and incentive.

Marcos Soares Lopes for your help, nice discussions and assistance with the data since my masters.

Topigs Norsvin and its researchers, especially Egbert, Marleen, Marcos, Claudia and Barbara, for the partnership, for providing the data, sharing knowledge and helping me every time I needed in the Netherlands. I learned a lot with you!

TABLE OF CONTENTS

ABSTRACT

MARQUES, Daniele Botelho Diniz, D.Sc., Universidade Federal de Viçosa, September, 2017. **Genetic parameters and genomic analysis of semen quality and fertility traits in pigs.** Advisor: Paulo Sávio Lopes. Co-advisors: Simone Eliza Facioni Guimarães and Fabyano Fonseca e Silva.

The widespread use of artificial insemination (AI) has greatly contributed to the success of the pig industry by assisting and disseminating the genetic progress. Currently, young boars are selected for AI based on their breeding values for production traits and selecting boars for semen traits, such as volume, concentration, motility and morphology, and for low variation in semen quality and production is still not a common practice. This selection is important for better performance of boars at AI stations, whose objective is to maximize the number of insemination doses produced by each ejaculate. The estimation of genetic parameters and the quantification of genetic variation for semen traits and within-boar variation allow an analysis of whether these traits should be included in the breeding goal. Besides the estimation of genetic parameters for selection purposes, the interest in studying the molecular processes and genetic mechanisms affecting semen traits is increasing in recent years. Genome-wide association studies (GWAS) are commonly used to identify single nucleotide polymorphisms (SNPs) associated with quantitative trait loci (QTL) with major effect. The weighted single-step GWAS (WssGWAS) is a method that allows estimation of SNP effects using information from all genotyped, phenotyped and pedigree animals. Expanding the frontiers of reproduction studies in pigs, another important field to be explored in breeding programs is the boar fertility. Reproductive traits, such as gestation length (GL), total number of piglets born (TNB) and stillborn (SB) are some of the bottleneck traits for efficient pig production. Because of the low to moderate heritabilities for these traits, it is important to identify all factors influencing them and to include these factors in the genetic evaluation models. The service sire (boar which ejaculate dose was used to inseminate the sow) and ejaculate effects are two of those important factors that have the potential to improve the traditional models used in the genetic evaluations of reproductive traits. Among the elements controlling the litter size, the fertilization rate and prenatal survival rate might be influenced by the service sire due to genetic differences in the capacity of fertilization, which is related to sperm quality and/or the boar genetic contribution to viability of the embryo. In this context, my overall aims were 1) to estimate genetic parameters for semen quality and quantity traits, as well as for within-boar variation of these traits; 2) to identify

QTL regions and candidate genes associated with semen traits through a WssGWAS and, subsequently, to perform gene network analyses to investigate the biological processes shared by genes identified in different pig lines and 3) to estimate genetic parameters for service sire on reproductive traits GL, TNB and SB and to evaluate the inclusion of service sire and ejaculate effects in the genetic evaluation models of these traits. The results of this thesis showed moderate estimates of heritability and favorable genetic correlations between semen traits, indicating that boar selection for these traits could make reasonable genetic progress. In addition, relevant genetic variation was found for within-boar variability of these traits, revealing the possibility of selection of boars for reduced variation in semen quality and production. Results from WssGWAS pinpointed relevant QTL regions explaining high proportions of genetic variance (up to 10.8%) for semen traits in several pig chromosomes, confirming the assumption of genetic complexity of these traits. This identification was possible with low number of animals having both phenotypes and genotypes due to the appropriate choice of the method. Candidate genes *SCN8A, PTGS2, PLA2G4A, DNAI2, IQCG*, LOC102167830, *NME5, AZIN2, SPATA7, METTL3* and *HPGDS* were identified associated with semen traits in the QTL regions identified for the pig lines evaluated. The gene network analysis showed candidate genes found for different pig lines sharing biological pathways that occur in mammalian testes. Regarding boar fertility, the results showed that there is genetic variation due to service sire effect on GL, TNB and SB; and the model with inclusion of permanent environmental and genetic effects due to service sire, in addition to ejaculate effect, showed the best fit to the data. This thesis resulted in important and innovative scientific information on male reproduction field in pigs, which will contribute to increase the still scarce knowledge about genetic selection and genomic architecture of boar semen quality and fertility traits.

RESUMO

MARQUES, Daniele Botelho Diniz, D.Sc., Universidade Federal de Viçosa, setembro de 2017. **Parâmetros genéticos e análise genômica de características de qualidade de sêmen e fertilidade em suínos.** Orientador: Paulo Sávio Lopes. Coorientadores: Simone Eliza Facioni Guimarães e Fabyano Fonseca e Silva.

O uso generalizado da inseminação artificial (IA) contribuiu grandemente para o sucesso da indústria de suínos, por meio do auxílio e disseminação do progresso genético. Atualmente, reprodutores suínos jovens são selecionados para IA com base em seus valores genéticos para características de produção e, a seleção de reprodutores para características de sêmen, como volume, concentração, motilidade e morfologia, bem como para menor variação intra-reprodutor na sua produção e qualidade, ainda não é uma prática comum. Esta seleção é importante para melhorar o desempenho dos reprodutores nas estações de IA, cujo objetivo é maximizar o número de doses inseminantes produzidas por cada ejaculado. A estimação de parâmetros genéticos e quantificação da variação genética para características de sêmen e para variação intra-reprodutor permitem analisar se essas características devem ser incluídas nos objetivos do melhoramento. Além da estimação de parâmetros genéticos para fins de seleção, o interesse em estudar os processos moleculares e os mecanismos genéticos que afetam as características de sêmen está aumentando nos últimos anos. Os estudos de associação genômica ampla (GWAS) são comumente usados para identificar polimorfismos de base única (SNPs) associados a *loci* de características quantitativas (QTL) com maiores efeitos. O GWAS em passo único ponderado (WssGWAS) é um método que permite a estimação de efeitos de SNP utilizando informações de todos os animais genotipados, fenotipados e com pedigree na população. Expandindo as fronteiras dos estudos de reprodução em suínos, outro campo importante a ser explorado em programas de melhoramento é a fertilidade dos reprodutores. As características reprodutivas, como a duração da gestação (GL), o número total de leitões nascidos (TNB) e nascidos mortos (SB) são algumas características-chave para a produção eficiente de suínos. Devido às baixas ou moderadas herdabilidades para essas características, é importante identificar todos os fatores que as influenciam e incluir esses fatores nos modelos de avaliação genética. Os efeitos do reprodutor cujo ejaculado foi utilizado para inseminar a matriz e do ejaculado são dois desses fatores importantes que têm o potencial de melhorar os modelos tradicionais utilizados nas avaliações genéticas das características reprodutivas. Dentre os elementos que controlam o tamanho da leitegada, as taxas de fertilização e de sobrevivência pré-natal podem ser

influenciadas pelo reprodutor, devido às diferenças genéticas na capacidade de fertilização relacionadas à qualidade do sêmen e/ou à contribuição genética do reprodutor para a viabilidade do embrião. Nesse contexto, os objetivos gerais com este estudo foram 1) estimar os parâmetros genéticos para qualidade e quantidade de sêmen, bem como para a variação intra-reprodutor para essas características; 2) identificar regiões de QTL e genes candidatos associados a características de sêmen por meio do WssGWAS e, subsequentemente, realizar análises de redes gênicas para investigar os processos biológicos compartilhados por genes identificados em diferentes linhas de suínos e 3) estimar parâmetros genéticos para o efeito do reprodutor na GL, TNB e SB e avaliar a inclusão dos efeitos do reprodutor e do ejaculado nos modelos de avaliação genética dessas características. Os resultados desta tese mostraram estimativas moderadas de herdabilidade e correlações genéticas favoráveis entre características de sêmen, indicando que a seleção de reprodutores para essas características pode resultar em razoável progresso genético. Além disso, variação genética relevante foi encontrada para a variabilidade intra-reprodutor para essas características, revelando a possibilidade de seleção de reprodutores para uma menor variação na qualidade e produção de sêmen. Os resultados do WssGWAS apontaram regiões relevantes de QTL que explicaram grandes proporções da variância genética (até 10,8%) para as características de sêmen em vários cromossomos suínos, confirmando a suposição de complexidade genética dessas características. Esta identificação foi possível com o baixo número de animais com fenótipos e genótipos, devido à escolha apropriada do método. Os genes candidatos *SCN8A*, *PTGS2*, *PLA2G4A*, *DNAI2*, *IQCG*, *LOC102167830*, *NME5*, *AZIN2*, *SPATA7*, *METTL3* e *HPGDS* foram identificados associados às características de sêmen nas regiões de QTL identificadas para as linhas de suínos avaliadas. A análise de redes gênicas mostrou genes candidatos encontrados para diferentes linhas de suínos compartilhando caminhos biológicos que ocorrem nos testículos de mamíferos. No que diz respeito à fertilidade do reprodutor, os resultados mostraram que há variação genética devido ao efeito do reprodutor em GL, TNB e SB; e o modelo com inclusão de efeitos de ambiente permanente e genéticos do reprodutor, além do efeito do ejaculado, mostrou o melhor ajuste para os dados. Esta tese resultou em informações científicas importantes e inovadoras na área de reprodução em machos, o que contribuirá para aumentar o conhecimento ainda escasso sobre a seleção genética e a arquitetura genômica de características de qualidade de sêmen e de fertilidade em reprodutores suínos.

CHAPTER 1

GENERAL INTRODUCTION

Male reproduction

The male gonads (testes) have the role of producing spermatozoa and hormones, mainly testosterone. The epididymis provides the ideal environment for spermatozoa maturation and storage before ejaculation. The boar presents well developed seminal vesicles and bulbourethral glands, which are responsible for production of a large semen volume compared to others livestock species (Bracket, 2006).

The process of producing spermatozoa, called spermatogenesis, occurs in the seminiferous tubules of mammalian testes in three steps: mitotic phase, meiotic phase, and spermiogenesis (Li et al., 2014). In the first step (mitosis), spermatogonias produce primary spermatocytes, which enter the first stage of meiosis (meiosis I), producing secondary spermatocytes. Then haploid round spermatids are generated after the second step of meiosis (meiosis II). In the last phase, spermiogenesis, the spermatids undergo morphological transformations, getting spermatozoa shape. Then, the new pre-formed spermatozoa go through epididymis to maturate and acquire motility (Li et al., 2014). The spermatogenesis process is complex and requires coordination among different cell types (germ cells, Sertoli cells, Leydig cells) (Sarkar et al., 2016) and genes. The Sertoli cells have a crucial role in the secretion of hormones and proteins required for normal spermatogenesis and the Leydig cells produce testosterone (Bracket, 2006). In pigs, the spermatogenesis cycle occur in 9 days and the period of time necessary to spermatogonias become ejaculated spermatozoa ranges between 50 to 60 days (Bracket, 2006).

Sperm production begins during puberty and males become able to have semen collected and spread when sexual maturity is reached. Boars reach puberty in about 5 to 6 months age and maturity about 8 to 9 months age (Bracket, 2006). The main components of spermatozoa are shown in Figure 1. Covering large part of the sperm head is the acrosome, a vesicle containing enzymes crucial in fertilization process. In the center of all extension of sperm tail is the axoneme, a microtubular structure displaying nine pair of microtubule surrounding a single pair of microtubule (McLachlan et al., 2012). Interacting with the

microtubules are the dyneins, ATPases that generate driving force for sperm motility (Witman, 1992).

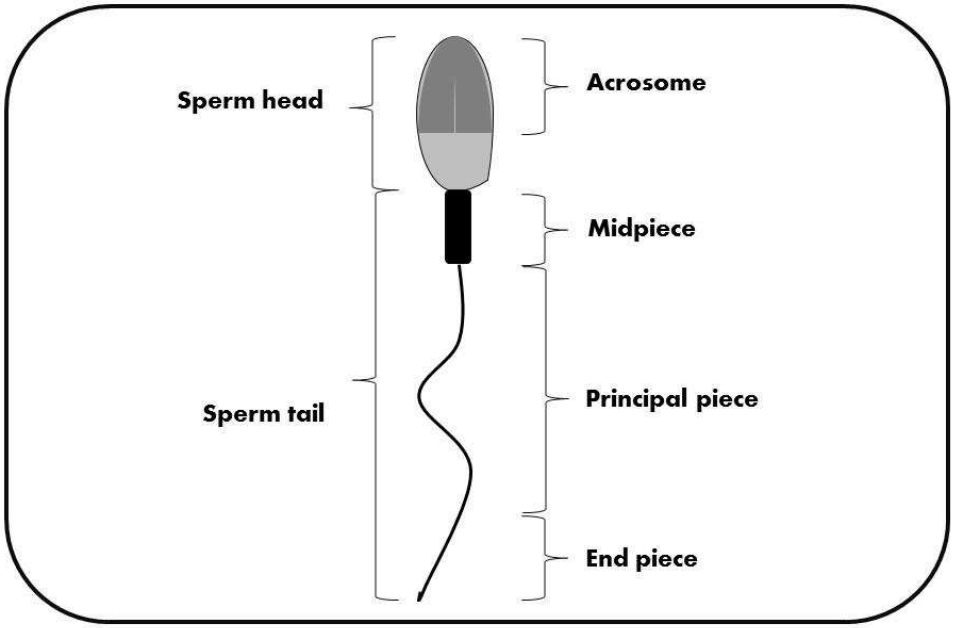

Figure 1. Main components of mammalian spermatozoa.

Semen traits and their assessment

Artificial insemination (AI) is widely used in pig industry. According to Knox (2016), efficiency and profitability in swine commercial production are linked to the use of AI, showing good results in pig fertility in relation to farrowing rates (FR) and total number of piglets born (TNB). In addition, AI allows accelerated genetic progress (Knox, 2016) by increasing the frequency of favorable alleles in the population. The main goal of AI industry is to maximize the number of high quality semen doses produced by each ejaculate of a boar. For that, the assessment of boar semen quality and quantity is a routine practice in AI stations and can be done by recording and evaluating sperm and ejaculate traits, such as motility, morphology, concentration, volume and number of sperm cells per ejaculate.

The sperm motility (proportion of moving sperm cells in the ejaculate) and the progressive motility (proportion of sperm cells moving in a straight line) are the most widely used traits in sperm quality evaluations and have been related to sperm fertilizing ability

(Gadea, 2005; Broekhuijse et al., 2012). Motility is acquired during sperm maturation in epididymis. In the course of epididymis transit, the volume of sperm cell changes, the nucleus condenses and the acrosome increases its maturation. The cytoplasmic droplet moves to distal position of sperm tail until its dissipation before ejaculation (Bracket, 2006).

Another important semen trait, essential for high semen quality, is the proportion of morphological abnormalities. These abnormalities include abnormal or loose acrosome, loose head or abnormal head shape, abnormal or bent tail and proximal and distal cytoplasmic droplets. In high proportions, these defects are able to reduce boar fertility (Alm et al., 2006; Bracket, 2006).

Semen traits motility, progressive motility, concentration and morphology can be evaluated by a technician with a microscope or by automatic systems, as the Computer assisted sperm analysis (CASA). The microscopic evaluation has been successful and important for many years for ejaculate quality control. The introduction of the CASA system has revolutionized the evaluation of ejaculates in AI studs (Broekhuijse et al., 2011; Knox 2016). The system is composed by a hardware and software used to visualize and analyze successive images of sperm, providing precise information on the movement of individual cells (Amann and Katz, 2004). Nowadays, a wide range of CASA systems are commercially available and, together with other softwares, it is also possible to calculate dilution rates for production of AI doses (Knox, 2016).

Genetic selection for semen traits in pigs

Traditional boar breeding programs have based their selection mainly on economically important production traits, such as growth rate, feed efficiency and lean yield, neglecting the potential benefits of selection for improved boar semen traits on fertility (Bobinson and Buhr, 2005; Diniz et al., 2014). Schulze et al. (2014) showed that 47.3% of young boars genetically proven for production traits were excluded from AI center because some semen traits did not surpass quality thresholds. Commercially important production traits should not be the unique criteria for boar genetic selection (Robinson and Buhr, 2005) because this may result in reduced semen production because of unfavorable genetic correlations between some production and semen traits (Oh et al., 2006a).

The inclusion of semen traits in breeding goals will depend on the heritabilities of the traits, the genetic correlations between semen and other target traits included in the selection

index and the objectives of the breeding program. Some advantages of genetic selection of boars for semen traits are: improvement of the proportion of boars that will meet the semen requirements at the AI stations; reduction of the discard of ejaculates that do not reach the AI threshold criteria; and increased genetic gain due to possible reduction in generation interval, since genetic selection of boars based on breeding values for semen traits will give a partial indication (because of moderate heritabilities) of the boars' performance before reach sexual maturity. The actual system in most countries, in which semen quality and production are analyzed only in AI centers, slows down the speed of genetic improvement (Rodriguez et al., 2017).

Some studies have been performed to estimate genetic parameters for semen traits in pigs (Smital et al., 2005; Oh et al., 2006b; Oh and See, 2008; Wolf, 2009a,b; Wolf and Smital, 2009a,b; Wolf, 2010). The authors described low to moderate heritabilities for these traits (0.05 ± 0.01 to 0.38 ± 0.04 for progressive motility; 0.10 ± 0.02 to 0.79 ± 0.09 for number of sperm cells per ejaculate and 0.04 ± 0.02 to 0.34 ± 0.03 for total morphological abnormalities). Genetic correlations estimated in those studies between progressive motility and morphological abnormalities were moderate and favorable: ranging between -0.34 ± 0.052 to -0.93 ± 0.088. Genetic correlations between progressive motility and number of sperm cells per ejaculate and between morphological abnormalities and number of sperm cells estimated by Smital et al. (2005) were -0.06 ± 0.058 and 0.14 ± 0.057, respectively.

The different results found for genetic parameters in these studies can be explained by different methodologies applied, diverse amount of data, different effects included in the statistical models, different breeds and crosses used and the divergent methods to measure semen traits (CASA system or microscope), which make comparison between studies more difficult.

Genetic selection for within-boar variation of semen traits

The within-boar variation of semen traits can be defined as the differences between records of semen traits of different ejaculates from the same boar. Boars may present a high variation in semen production and quality among ejaculates, which is a concern of the AI industry because leads to the exclusion of ejaculates that do not reach the threshold for the production of insemination doses, and, consequently, induces to economic loss. In the pig breeding sector, genetic selection for this variation is not practiced and in literature, studies

evaluating the possibility of boars' selection based on variation of semen traits are not performed. Therefore, the estimation of genetic parameters for within-boar variation is important in order to evaluate and optimize the selection of boars that present less fluctuation in semen quality and production. The within-boar variation can be evaluated, for example, by the standard deviation (in addition to the mean) or coefficient of variation of traits' records. In literature, the majority of studies regarding variation describes the estimation of genetic parameters for mean and standard deviation or coefficient of variation of production traits, such as within-litter variation of piglet weights (Hermesch et al., 2001; Damgaard et al., 2003; Wolf et al., 2008) and variation of somatic cell score in dairy cattle (Urioste et al., 2010; Wijga et al., 2012).

Genomic regions associated with semen traits in pigs

In recent years, with the fast advances in high-throughput genotyping and in molecular techniques, the interest in studying the molecular processes affecting boar semen traits has increased. Genes and markers associated with pig semen traits in different chromosomes have been described in literature (Xing et al., 2009; Sironen et al., 2010; Kaewmala et al., 2011; Gunawan et al., 2011; Gunawan et al., 2012; Diniz et al., 2014; Zhao et al., 2016). Genome-wide significant quantitative trait loci (QTL) were identified for semen pH, volume and ejaculation times (Xing et al., 2009). In relation to candidate genes, the following associations have been reported: *HECW2* gene with sperm acrosome defect (Sironen et al., 2010); *CD9* gene with sperm development (Kaewmala et al., 2011); *ESR1* (Guanawan et al., 2011) and *ERS2* (Guanawan et al., 2012) genes with spermatogenesis; *MTFMT* gene with sperm motility (Diniz et al., 2014) and *APN, TEP1, PARP2, SPINK1* and *PDE1C* genes with 13 boar reproductive traits related to testes, semen and ejaculation (Zhao et al., 2016).

Boar fertility

Reproductive traits, such as gestation length (GL), TNB and stillborn (SB), can be influenced by the service sire, which is the boar whose semen was used to inseminate the sow, and by the ejaculate effect. Traditionally, breeding programs ignore the paternal side or include the service sire in the genetic evaluation models of reproductive traits just as a non-genetic random effect. However, according to Van der Lende et al. (1999), the influence of

the service sire may be due to the genetic contribution of the boar to viability of the embryo. The main reason for breeding programs do not include the effect of service sire in genetic evaluation models for reproductive traits may be the low genetic variance and heritabilities for service sire based on these traits (Van der Lende et al., 1999; Serenius et al., 2003; Hamann et al., 2004; Su et al., 2007; Wolf and Wolfová et al., 2012).

In literature, there are some reports on genetic parameter estimations for service sire on reproductive traits and the conclusions are controversy. Van der Lende et al. (1999) evaluated the service sire effect on TNB and number of piglets born alive (NBA) in 36,708 litters from 1,044 service sires from different pure lines and one cross. Heritability estimates for service sire effect on TNB and NBA ranged from 0.002 to 0.04 and the authors concluded that genetic selection for service sire effect would not improve litter size in pigs. Serenius et al. (2003) analyzed litter records of 6,514 Finnish Large White and 9,154 Landrace pigs. Heritability estimates for service sire effect on TNB ranged from 0.008 to 0.012. The authors concluded that besides the low effect of service sire, it should be considered in the statistical model as a random environmental effect for culling of boars with poor piglet production. Wolf and Wolfová (2012) evaluated the inclusion of service sire in genetic evaluation models for TNB, NBA and number of piglets weaned. The authors estimated heritabilities for service sire on TNB, NBA and number of piglets weaned in the range from 0.02 to 0.03 and concluded that models without service sire effect or models including service sire just as random effect are proposed for genetic evaluation of litter size traits. Su et al. (2007) reported service sire heritability estimates on TNB ranging from 0.02 to 0.03. They concluded that it is reasonable to include the service sire effect in the model for genetic evaluation of litter size.

Differences in semen quality are also reported as responsible for divergent boar fertility, although some authors stated that the semen analyses are not the best predictor of boar fertility (Gadea et al., 2004; Ruiz-Sánchez et al., 2006). Broekhuijse et al. (2012) analyzed data from 45,532 boar ejaculates and showed that 5.3% of the total variation in FR and 5.9% of the total variation in TNB were explained by the direct boar effect (effect of individual boar and semen related traits). Sperm motility explained 9% of the boar and semen related variation in FR and 10% of the boar and semen related variation in TNB. Wolf (2010) evaluated litter size data from 28,485 Czech Large White and 10,410 Czech Landrace sows and reported a trend in increasing litter size associated with slight decrease in the total number

of sperm and in the number of functional sperm, especially in the Large White breed. Gadea et al. (2004) analyzed data of FR and TNB from 1,818 sows and 273 ejaculates from 57 boars to evaluate the fertility predictive value of different sperm traits. The authors concluded that it was not possible to detect fertility differences associated with semen traits because of the high number of sperm cells per ejaculate dose and the high quality of the semen used in AI.

This thesis: objectives

My aim with this thesis was to estimate the genetic parameters for semen quality and quantity traits in pigs and for within-boar variation of semen traits to evaluate their inclusion in breeding goals. In addition, my objective was also to report genomic regions and candidate genes associated with semen traits to clarify the genetic control of such traits. I also evaluated the effect of service sire and ejaculate in reproductive traits of pigs to verify the paternal effect on pig fertility.

In Chapter 2, heritabilities and genetic correlations were estimated for four semen traits using multiple-trait analysis in five different commercial pig lines. Additionally, genetic parameters were also estimated for within-boar variation for those four semen traits and the capacity of variation traits to account for uniformity in semen quality was accessed. In Chapter 3, a weighted single-step GWAS was performed for a Landrace-based and a Large White-based pig lines to find QTL regions associated to the semen traits motility, progressive motility, number of sperm cells per ejaculate and morphological defects. In addition, candidate genes associated to one or more semen traits were selected within the QTL regions identified in GWAS and four gene network analyses for biological processes shared by candidate genes in different lines were applied. In chapter 4, I estimated genetic parameters for service sire on reproductive traits GL, TNB and SB and evaluated the inclusion of genetic service sire and ejaculate effects in the genetic evaluation models of those reproductive traits in a Large White-based dam line. In Chapter 5, I presented an overall vision of pig breeding schemes, showed the importance of AI in assisting the genetic progress in pig industry in addition to the main aspects of boar fertility and semen traits and discussed the results of the thesis in a broader context.

References

ALM, K.; PELTONIEMI, O. A.; KOSKINEN E.; ANDERSSON, M. Porcine field fertility with two different insemination doses and the effect of sperm morphology. **Reproduction in Domestic Animals**, v.41, p.210-213. 2006.

AMANN, R. P.; KATZ, D. F. Reflections on CASA after 25 years. **Journal of Andrology.** v. 25(3), p.317–325, 2004.

BRACKETT, B. G. Reprodução em Mamíferos do Sexo Masculino. **In: REECE, W. O. Dukes Fisiologia dos Animais Domésticos**, Rio de Janeiro: Guanabara Koogan, p.638, 2006.

BROEKHUIJSE, M. L. W. J.; FEITSMA, H.; GADELLA, B. M. Additional value of computer assisted semen analysis (CASA) compared to conventional motility assessments in pig artificial insemination. **Theriogenology,** v.76, p.1473–1486, 2011.

BROEKHUIJSE, M. L. W. J.; ŠOŠTARIĆ E.; FEITSMA, H.; GADELLA, B. M. Application of computer-assisted semen analysis to explain variations in pig fertility. **Journal of Animal Science,** v. 90(3), p.779–789, 2012.

DAMGAARD, L. H.; RYDHMER, L.; LØVENDAHL, P.; GRANDINSON, K. Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. **Journal of Animal Science,** v.81, p.604-610, 2003.

DINIZ, D. B.; LOPES, M. S.; BROEKHUIJSE, M. L. W. J.; LOPES, P. S.; HARLIZIUS, B. GUIMARÃES, S. E. F.; DUIJVESTEIJN, N.; KNOL, E. F.; SILVA, F. F. A genome-wide association study reveals a novel candidate gene for sperm motility in pigs. **Animal Reproduction Science**, v. 151. P. 201-207, 2014.

GADEA J. Sperm factors related to in vitro and in vivo porcine fertility. **Theriogenology**, v. 63, p.431-444, 2005.

GADEA, J.; SELLÉS, E.; MARCO, M. The predictive value of porcine seminal parameters on fertility outcome under commercial conditions. **Reproduction in Domestic Animals,** v.39, p.303-308, 2004.

GUNAWAN, A.; KAEWMALA, K.; UDDIN, M. J.; CINAR, M. U.; TESFAYE, D.; PHATSARA, C.; et al. Association study and expression analysis of porcine ESR1 as a

candidate gene for boar fertility and sperm quality. **Animal Reproduction Science**, v.128, p.11-21, 2011.

GUNAWAN, A.; CINAR, M. U.; UDDIN, M. J.; KAEWMALA, K.; TESFAYE, D.; PHATSARA, C.; et al. Investigation on association and expression of ESR2 as a candidate gene for boar sperm quality and fertility. **Reproduction in Domestic Animals**, v. 47, p.782-790, 2012.

HAMANN, H.; STEINHEUER, R.; DISTL, O. Estimation of genetic parameters for litter size as a sow and boar trait in German herdbook Landrace and Pietrain swine. **Livestock Production Science**, v. 85, p.201–207, 2004.

HERMESCH, S.; LUXFORD, B. G.; GRASER, H. -U. Genetic parameters for piglet mortality, within litter variation of birth weight, litter size and litter birth weight. **Proc. Association for the Advancement of Animal Breeding and Genetics.** p 211-214. 2001.

KAEWMALA K.; UDDIN, M. J.; CINAR, M. U.; GROSSE-BRINKHAUS, C.; JONAS, E.; TESFAYE, D.; et al. Association study and expression analysis of CD9 as candidate gene for boar sperm quality and fertility traits. **Animal Reproduction Science**, v.125, p.170-179, 2011.

KNOX, R. V. Artificial insemination in pigs today. **Theriogenology,** v. 85, p.83–93, 2016.

LI, R. -K.; TAN, J. -L.; CHEN, L. -T.; FENG, J. -S.; LIANG, W. -X.,; GUO, X. -J.; LIU, P.; CHEN, Z.; SHA, J. -H.; WANG, Y. -F. Iqcg is essential for sperm flagellum formation in mice. **PloS One**, v.9, p.e98053, 2014.

MCLACHLAN, R. I.; ISHIKAWA, T.; OSIANLIS, T.; ROBINSON, P.; MERRINER, D. J.; HEALY, D.; et al. Normal live birth after testicular sperm extraction and intracytoplasmic sperm injection in variant primary ciliary dyskinesia with completely immotile sperm and structurally abnormal sperm tails. **Fertility and Sterility**, v. 97, p.313-318, 2012.

OH, S. H.; SEE, M. T.; LONG, T. E.; GALVIN, J. M. Estimates of genetic correlations between production and semen traits in boar. **Asian-Australasian Journal of Animal Sciences,** v.19, p.160-164, 2006a.

OH, S. H.; SEE, M. T.; LONG, T. E.; GALVIN, J. M. Genetic parameters for various random regression models to describe total sperm cells per ejaculate over the reproductive lifetime of boars. **Journal of Animal Science**, v.84, p.538-545, 2006b.

OH, S. H.; SEE, M. T. Comparison of genetic parameter estimates of total sperm cells of boars between random regression and multiple trait animal models. **Asian-Australasian Journal of Animal Sciences,** v.21. p.923-927, 2008.

ROBINSON, J. A. B.; BUHR, M. M. Impact of genetic selection on management of boar replacement. **Theriogenology,** v. 63(2), p.668–678, 2005.

RODRIGUEZ, A. L.; SOOM, A. V.;. ARSENAKIS, I.; MAES, D. Boar management and semen handling factors affect the quality of boar extended semen. **Porcine Health Management**, v.3, p.15, 2017.

RUIZ-SÁNCHEZ, A. L.; O'DONOGHUE, R.; NOVAK, S.; DYCK, M. K.; COSGROVE, J. R.; DIXON, W. T.; FOXCROFT, G. R. The predictive value of routine semen evaluation and IVF technology for determining relative boar fertility. **Theriogenology,** v.66, p.736–748, 2006.

SARKAR, H.; ARYA, S.; RAI, U.; MAJUMDAR, S. S. A study of differential expression of testicular genes in various reproductive phases of *hemidactylus flaviviridis* (wall lizard) to derive their association with onset of spermatogenesis and its relevance to mammals. **PloS One,** v.11, p.e0151150, 2016.

SCHULZE, M.; BUDER, S.; RÜDIGER, K.; BEYERBACH, M.; WABERSKI, D. Influences on semen traits used for selection of young AI boars. **Animal Reproduction Science**, v.148, p.164-170, 2014.

SERENIUS, T.; SEVÓN-AIMONEN, M. -L.; MANTYSAARI, E. A. Effect of service sire and validity of repeatability model in litter size and farrowing interval of Finnish Landrace and Large White populations. **Livestock Production Science**, v.81, p.213–222, 2003.

SIRONEN, A.; UIMARI, P.; NAGY, S.; PAKU, S.; ANDERSSON, M.; VILKKI, J. Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. **BMC Genomics**, v.11, p.699, 2010.

SMITAL, J.; WOLF, J.; DE SOUSA, L. L. Estimation of genetic parameters of semen characteristics and reproductive traits in AI boars. **Animal Reproduction Science**, v.86. p.119-130, 2005.

SU G.; LUND, M. S.; SORENSEN, D. Selection for litter size at day five to improve litter size at weaning and piglet survival rate. **Journal of Animal Science**, v.85, p.1385–1392, 2007.

URIOSTE, J. I.; FRANZÉN, J.; STRANDBERG, E. Phenotypic and genetic characterization of novel somatic cell count traits from weekly or monthly observations. **Journal of Dairy Science**, v. 93, p.5930-5941, 2010.

VAN DER LENDE, T.; WILLEMSEN, M. H. A.; VAN ARENDONK, J. A. M.; VAN HAANDEL, E. B. P. G. Genetic analysis of the service sire effect on litter size in swine. **Livestock Production Science**, v.58. p.91–94, 1999.

WIJGA, S.; BASTIAANSEN, J. W. M.; WALL, E.; STRANDBERG, E.; DE HAAS, Y.; GIBLIN, L. BOVENHUIS, H. Genomic associations with somatic cell score in first-lactation Holstein cows. **Journal of Dairy Science**, v.95, p.899-908, 2012.

WITMAN G.B. Axonemal dyneins. **Current Opinion in Cell Biology**, v. 4, p.74-79, 1992.

WOLF, J. Genetic correlations between production and semen traits in pig. **Animal,** v. 3, p.1094-1099, 2009a.

WOLF, J. Genetic parameters for semen traits in AI boars estimated from data on individual ejaculates. **Reproduction in Domestic Animals**, v. 44, p.338-344, 2009b.

WOLF, J. Heritabilities and genetic correlations for litter size and semen traits in Czech Large White and Landrace pigs. **Journal of Animal Science**, v. 88, p.2893-2903, 2010.

WOLF, J.; SMITAL, J. Effects in genetic evaluation for semen traits in Czech Large White and Czech Landrace boars. **Czech Journal of Animal Science**, v.54, p.349-358, 2009a.

WOLF, J.; SMITAL, J. Quantification of factors affecting semen traits in artificial insemination boars from animal model analyses. **Journal of Animal Science,** v.87, p.1620-1627, 2009b.

WOLF J.; WOLFOVÁ, M. Effect of service sire on litter size traits in Czech Large White and Landrace pigs. **Czech Journal of Animal Science**, v.57, p.220-230, 2012.

WOLF, J.; ŽÁKOVÁ, E.; GROENEVELD, E. Within-litter variation of birth weight in hyperprolific Czech Large White sows and its relation to litter size traits, stillborn piglets and losses until weaning. **Livestock Science**, v.115, p.195-205, 2008.

XING, Y.; REN, J.; REN, D.; GUO, Y.; WU, Y.; YANG, G.; et al. A whole genome scanning for quantitative trait loci on traits related to sperm quality and ejaculation in pigs. **Animal Reproduction Science,** v.114, p.210-218, 2009.

ZHAO, X.; ZHAO, K.; REN, J.; ZHANG, F.; JIANG, C.; HONG, Y.; et al. An imputation-based genome-wide association study on traits related to male reproduction in a White Duroc x Erhualian F2 population. **Animal Science Journal,** v. 87(5), p.646-54, 2016.

CHAPTER 2

Genetic parameters for semen quality and quantity traits in five pig lines[1]

D.B.D. Marques,* M.S. Lopes,†‡ M.L.W.J. Broekhuijse,† S.E.F.Guimarães,* E.F. Knol,† J.W.M. Bastiaansen,§ F.F. Silva,* P.S. Lopes*

*Universidade Federal de Viçosa, Animal Science Department, 36570-000, Viçosa-MG, Brazil; †Topigs Norsvin Research Center B.V., P.O. Box 43, 6640 AA Beuningen, the Netherlands; ‡Topigs Norsvin, 80420-210, Curitiba-PR, Brazil; and §Wageningen University & Research, Animal Breeding and Genomics, P.O. Box 338 6700 AH, Wageningen, the Netherlands

[1]The dataset used for these analyses was based on the results of the ejaculates of boars located at the centres of AIM Varkens KI Nederland (Vught, the Netherlands). The authors thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the financial support. We are also grateful to Topigs Norsvin for providing the data.

Journal of Animal Science
Received April 30, 2017
Accepted August 15, 2017
doi:10.2527/jas2017.1683.

ABSTRACT: We aimed to estimate genetic parameters for semen quality and quantity traits as well as for within-boar variation of these traits to evaluate their inclusion in breeding goals. Genetic parameters were estimated within line using a multiple-trait (4 × 4) repeatability animal model fitted for 5 pig lines, considering 4 semen traits: sperm motility (MOT), sperm progressive motility (PROMOT), log-transformed number of sperm cells per ejaculate (lnN_{cells}), and total morphological abnormalities (ABN). The within-boar variation of these traits was analyzed based on a multiple-trait (2 × 2) approach for SD and average (AVG) and a single-trait analysis for CV. The average heritabilities across the 5 lines estimated by multiple-trait analysis were 0.18 ± 0.07 (MOT), 0.22 ± 0.08 (PROMOT), 0.16 ± 0.04 (lnN_{cells}), and 0.20 ± 0.04 (ABN). The average genetic correlations were favorable between MOT and PROMOT (0.86 ± 0.10), between MOT and ABN (−0.66 ± 0.25), and between PROMOT and ABN (−0.65 ± 0.25). As determined by within-boar variation analysis, AVG exhibited the greatest heritabilities followed by SD and CV, respectively, for the traits MOT and ABN. For PROMOT, average SD heritability was lower than CV heritability, whereas for lnN_{cells}, they were the same. The average genetic correlations between AVG and SD were favorable for MOT (−0.60 ± 0.13), PROMOT (−0.79 ± 0.14), and ABN (0.78 ± 0.17). The moderate heritabilities indicate the possibility of effective selection of boars based on semen traits. Average and SD are proposed as appropriate traits for selection regarding uniformity.

Key words: multiple-trait analysis, pig, selection, semen, variation analysis

INTRODUCTION

The widespread use of AI has greatly contributed to the success of the pig industry by assisting and disseminating the genetic progress. Currently, young boars are selected for AI based on their breeding values regarding production traits. However, the economic profitability of an AI station also depends on the evaluation of semen quantity and quality traits, such as volume, concentration, motility, and morphology (Wolf, 2009b; Broekhuijse et al., 2011a). The efficiency of a selection program depends on knowledge of the genetic covariance structure of target traits (Smital et al., 2005). The estimation of genetic correlations between these traits is an important step for determining their inclusion in the selection index and for calculating the correlated selection response. To date, however, only a few studies have been performed to estimate the genetic parameters for boar semen traits (Smital et al., 2005; Oh et al., 2006b; Wolf, 2009a,b, 2010; Wolf and Smital, 2009a,b). A concern of the AI industry is that some boars exhibit a high variation in semen quality and quantity traits over the course of a year. This variation leads to the exclusion of ejaculates that do not reach the threshold for the production of insemination doses, leading to economic loss. Therefore, in addition to the importance of estimating the genetic parameters for these traits, it is also important to evaluate the within-boar variation and to optimize the selection of boars for both the mean and variation of these traits. The within-boar variation can be defined as the differences between records of semen traits of different ejaculates from the same boar, measured as SD or CV. In this study, we aimed to estimate genetic parameters for semen quality and quantity traits as well as for within-boar variation based on data collected from 5 pig lines. The ability of variation traits to quantify uniformity was also assessed.

MATERIALS AND METHODS

Ethics Statement

The data used for this study were obtained as part of routine data recording in a commercial breeding program. Data recording and sample collection were conducted in strict accordance with Dutch law on the protection of animals (Gezondheids- en welzijnswet voor dieren).

Animals and Semen Traits

Data from 5 pig lines were available: a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds (**HY**), a Pietrainbased line (**PT**), a Duroc-based line (**DU**), a Large White–based line (**LW**), and a Landrace-based line (**LD**). The boars were managed according to standard protocols of the AI stations. Each boar was individually housed in barns with adequate ventilation and air movement and was maintained at an acceptable ambient temperature. Water was provided ad libitum, and the boars were fed 2 times a day. Procedures were conducted according to International Standards Organization (**ISO**) certifed protocols (ISO 9001:2008) in accordance with the Dutch Experiments on Animals Act of 1998 (AIM Varkens KI Nederland B.V., Vught, the Netherlands).

The following semen traits were analyzed: 1) quality traits, including sperm motility (**MOT**), which is the proportion of moving sperm cells in an ejaculate; sperm progressive motility (**PROMOT**), defined as the proportion of sperm cells moving in a straight line; and the percentage of total morphological abnormalities (**ABN**), which are sperm cells with morphological ab abnormalities, and 2) a quantity trait: the total number of sperm cells in the ejaculate ($\mathbf{N_{cells}}$; in 10^6 sperm cells).

Boar ejaculates were collected on a routine basis at the AI stations, and the semen was processed and analyzed according to the ISO (ISO 9001:2008) certified protocols in accordance with the Dutch Experiments on Animals Act of 1998 (described in detail by Broekhuijse et al., 2011b). The ejaculate samples evaluated for MOT and PROMOT were collected between January 2007 and October 2014 and were evaluated using the UltiMate CASA system (Hamilton Thorne Inc., Beverly, MA). Total number of sperm cells in the ejaculate was calculated as the product of the semen volume (mL) and concentration (106 mL−1, measured by the computer-assisted semen analysis [**CASA**] system), traits evaluated from ejaculates collected between January 2007 and October 2014. Total number of sperm cells in the ejaculate was not normally distributed and was therefore log-transformed (log-transformed number of sperm cells per ejaculate [$\mathbf{lnN_{cells}}$]). Ejaculates evaluated for ABN were collected between January 1997 and October 2014 and were microscopically analyzed by a technician. All semen traits were assessed on the day of semen collection (fresh).

The numbers of boars and ejaculates per line from the phenotypic data are summarized in Table 1. The synthetic line presented the largest number of ejaculates and boars (221,602 and 2,947, respectively), whereas LD had the smallest number of ejaculates (39,161) and LW had the smallest number of boars (866). Mean values for the traits were highly similar among the lines. The biggest differences were observed for ABN, with LD exhibiting a value of 14.37, which was lower than the other lines, which ranged from 15.85 to 19.27 (Table 2).

The age of the boar at the semen collection and the interval between 2 subsequent semen collections ranged from 8 to 72 mo and from 1 to 14 d, respectively, in all lines, except for PT, for which the age of the boar ranged from 9 to 72 mo. For HY, PT, DU, LW, and LD, the mean ages of the boars were 24.11 ± 10.78, 26.31 ± 11.97, 23.25 ± 10.12, 19.11 ± 8.03,

and 17.56 ± 6.28, respectively, and the intervals between 2 subsequent semen collections were 5.93 ± 1.89, 6.06 ± 2.08, 6.05 ± 2.01, 5.40 ± 2.45, and 5.37 ± 2.49, respectively.

For all lines, up to 5 generations of pedigree data were available. The pedigree data for HY, PT, DU, LW, and LD contained 6,646, 6,060, 3,933, 3,858, and 4,576 animals, respectively.

Table 1. Number of boars and ejaculates per line

Pig line [1]	Number of boars	Number of ejaculates	Ejaculates per boar (SD) [2]
HY	2,947	221,602	75.20 (56.18)
PT	2,544	200,620	78.86 (66.59)
DU	1,697	116,781	68.82 (50.72)
LW	866	43,455	50.18 (38.12)
LD	900	39,161	43.51 (36.37)

[1]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

[2]Average number of ejaculates per boar.

Table 2. Mean, SD, minimum, and maximum values[1] for semen quality and quantity traits[2] used in the multiple-trait analysis

Semen trait and pig line	Mean	SD	Minimum	Maximum
MOT				
HY	86.85	7.13	10.00	100.00
PT	87.65	6.68	10.00	100.00
DU	87.12	6.89	10.00	100.00
LW	86.50	7.13	10.00	100.00
LD	87.09	6.51	14.00	100.00

PROMOT				
HY	79.05	9.31	0.00	100.00
PT	79.50	8.53	0.00	100.00
DU	77.86	8.37	0.00	100.00
LW	78.57	8.30	0.00	100.00
LD	77.43	7.92	0.00	100.00
lnN_{cells}				
HY	11.23	0.39	7.87	12.74
PT	11.18	0.40	7.77	12.79
DU	11.11	0.43	7.49	13.20
LW	11.23	0.40	7.80	12.76
LD	11.12	0.39	8.69	12.55
ABN				
HY	16.63	13.01	1.00	100.00
PT	15.85	13.36	1.00	99.00
DU	17.91	12.73	1.00	99.00
LW	19.27	14.81	1.00	98.00
LD	14.37	12.56	1.00	99.00

[1]Mean, SD), minimum, and maximum values of sperm traits in the 5 evaluated lines: HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

[2]MOT = sperm motility (%); PROMOT = sperm progressive motility (%); lnN_{cells} = log-transformed number of sperm cells per ejaculate; ABN = total morphological abnormalities (%).

Genetic Parameters for Semen Traits from Multiple-Trait Analysis

Genetic parameters (narrow-sense heritabilities and genetic correlations) were estimated by multiple-trait analysis (4x4) using a repeatability animal model in ASReml 3.0 (Gilmour et al., 2009). Narrow-sense heritabilities (h^2) were calculated as $h^2 = V_a/(V_a + V_p + V_e)$, in which V_a, V_p, and V_e were the estimated additive genetic, permanent environmental and

residual variances, respectively. The analyses were performed within line applying the following model:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Za} + \mathbf{Wp} + \boldsymbol{\varepsilon}, \quad [1]$$

in which **y** is the vector of observations; **β** is the vector of fixed effects (combined effects of AI station, year and month of semen collection, the laboratory where the samples were analyzed, and the covariates interval between 2 subsequent semen collections in days and the age of the boar in months at the semen collection); **a** is the vector of random additive genetic effects; **p** is the vector of random permanent environmental effects; **ε** is the vector of random residuals; and **X**, **Z,** and **W** are the incidence matrices of **β, a,** and **p**, respectively.

It was assumed that $\mathbf{a} \sim \mathrm{N}(\mathbf{0}, \mathbf{A} \otimes \mathbf{G})$, $\mathbf{p} \sim \mathrm{N}(\mathbf{0}, \mathbf{I} \otimes \mathbf{P})$, and $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{I} \otimes \mathbf{R})$, where $\otimes$ is a direct product operation; **G**, **P,** and **R** represent covariance matrices for additive genetic, permanent environmental and residual effects, respectively; **A** is the pedigree-based relationship matrix; and **I** is an identity matrix.

Genetic Parameters for Within-Boar Variation of Semen Traits

To calculate the within-boar variation of semen traits, the observations were adjusted for all nongenetic effects before genetic analysis. The preadjusted data were obtained by adding the general mean for each trait to the predictions of genetic, permanent environmental, and residual effects obtained via single-trait analysis, as follows:

$$\mathbf{y^*} = \mathbf{y} - \mathbf{X\beta} = \mathbf{Za} + \mathbf{Wp} + \boldsymbol{\varepsilon},$$

in which **y*** is the vector of preadjusted data and **y**, **β, a**, **p**, **ε, X**, **Z,** and **W** are as previously defined.

Preadjustment was necessary because the fixed effects and covariates varied among ejaculates of the same animal and could not be fitted in a simple linear mixed model. Only animals with ≥10 ejaculates, which was considered the minimum number of observations sufficient to calculate the within-boar variation of the studied traits (Wijga et al., 2012), were included in this analysis,

Two traits were investigated to evaluate the within-boar variation of semen traits: 1) the average (**AVG**) and SD in a multiple trait analysis (2 x 2), where animals with greatest AVG are preferred, except for the trait ABN, and animals with lowest SD are preferred for all traits and 2) the CV in a single-trait analysis. The AVG and SD were calculated for each boar from the preadjusted observations. The CV combines the calculated AVG and SD into one trait:

$$\mathrm{CV} = (\mathrm{SD}/\mathrm{AVG}) * 100,$$

in which CV is expressed as a percentage.

Narrow-sense heritabilities for the variation traits were obtained from an additive animal model implemented in ASReml by applying a multiple-trait analysis (2 x 2) for AVG and SD and a single-trait analysis for CV. The h^2 were calculated as $V_a / (V_a + V_e)$, in which V_a and V_e were the estimated additive genetic and residual variances, respectively. Genetic correlations were estimated between AVG and SD for all lines and semen traits.

Single-trait analyses using the repeatability animal model [1] were performed for MOT, PROMOT, $\ln N_{cells}$ and ABN in this study, with the aim of comparing the variance

components and genetic parameters estimates from these single-trait analyses with those for AVG in the within-boar variation analysis.

RESULTS

Multiple-Trait Analysis (4 x 4)

Variance component estimates from the multiple-trait analysis (4 × 4) are provided in Supplemental File S1. Heritability estimates ranged from 0.08 ± 0.03 to 0.31 ± 0.05 across traits and lines (Table 3). In general, PROMOT and ABN showed the greatest heritabilities across the 5 lines, with average estimates of 0.22 ± 0.08 and 0.20 ± 0.04, respectively. For MOT and lnN_{cells}, heritability estimates were, on average, 0.18 ± 0.07 and 0.16 ± 0.04, respectively. Heritability estimates for LD were lower than the estimates for the other lines, except for lnN_{cells}, for which LD had a slightly greater estimate than LW (0.13 vs. 0.12; Table 3). High positive genetic correlations were observed between MOT and PROMOT for all lines (0.86 ± 0.10, on average, across the 5 lines). The genetic correlations were not different from 0 between MOT and lnN_{cells} and between PROMOT and lnN_{cells}, except for LW (genetic correlation [$\boldsymbol{r_g}$] = −0.67 ± 0.17 and r_g = −0.65 ± 0.15, respectively). Moderate and strong negative genetic correlations were found between MOT and ABN and between PROMOT and ABN, except for LD, for which the correlations were also negative but were weaker (r_g = −0.25 ± 0.26 and r_g = −0.27 ± 0.21, respectively). Genetic correlations between lnN_{cells} and ABN were mostly small and negligible, except for LW (r_g = 0.61 ± 0.17).

Table S1. Variance components (SE) from the multiple-trait analysis (4 x 4)

Semen trait[1]	Variance components[2]		
	σ_a^2	σ_{perm}^2	σ_ε^2
MOT			

HY[3]	12.13 (1.38)	10.49 (0.91)	30.46 (0.09)
PT	8.94 (1.69)	22.69 (1.44)	27.23 (0.09)
DU	8.29 (1.64)	12.84 (1.22)	28.20 (0.12)
LW	16.73 (4.05)	24.83 (3.13)	26.18 (0.18)
LD	3.38 (1.55)	19.52 (1.66)	21.69 (0.16)
PROMOT			
HY	20.58 (2.06)	12.35 (1.28)	40.92 (0.12)
PT	14.89 (2.36)	27.18 (1.90)	41.74 (0.13)
DU	15.31 (2.56)	16.88 (1.80)	40.62 (0.17)
LW	28.61 (5.73)	26.64 (4.05)	38.15 (0.26)
LD	7.05 (2.21)	21.78 (2.10)	35.63 (0.26)
lnN_{cells}			
HY	0.03 (3.80E-03)	0.04 (2.70E-03)	0.08 (3.00E-04)
PT	0.03 (4.60E-03)	0.05 (3.70E-03)	0.09 (3.00E-04)
DU	0.03 (4.40E-03)	0.03 (3.10E-03)	0.10 (4.00E-04)
LW	0.02 (4.30E-03)	0.04 (3.70E-03)	0.07 (5.00E-04)
LD	0.02 (5.10E-03)	0.04 (4.50E-03)	0.07 (5.00E-04)
ABN			
HY	33.61 (4.35)	43.17 (3.15)	98.72 (0.56)
PT	36.04 (6.68)	82.44 (5.60)	97.96 (0.57)
DU	45.98 (7.93)	55.87 (5.71)	91.75 (0.70)
LW	71.56 (18.00)	104.31 (14.09)	107.78 (1.39)
LD	25.96 (8.94)	76.78 (8.30)	73.89 (1.13)

[1]MOT = sperm motility; PROMOT = sperm progressive motility; lnN_{cells} = log-transformed number of cells per ejaculate; ABN = total morphological abnormalities.

[2]Variance components from the multiple trait analysis (4 x 4): σ_a^2 = additive genetic variance; σ_{perm}^2 = permanent environmental variance; σ_e^2 = residual variance.

[3]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

Table 3. Estimates of heritabilities (bold; SE) and genetic correlations (SE) between sperm **t**raits[1] from multiple-trait analysis

Semen trait	Pig line[2]	MOT	PROMOT	lnN_{cells}	ABN
MOT	HY	**0.23 (0.02)**			
	PT	**0.15 (0.03)**			
	DU	**0.17 (0.03)**			
	LW	**0.25 (0.05)**			
	LD	**0.08 (0.03)**			
PROMOT	HY	0.93 (0.01)	**0.28 (0.02)**		
	PT	0.81 (0.04)	**0.18 (0.03)**		
	DU	0.87 (0.03)	**0.21 (0.03)**		
	LW	0.96 (0.01)	**0.31 (0.05)**		
	LD	0.71 (0.12)	**0.11 (0.03)**		
lnN_{cells}	HY	0.01 (0.08)	0.03 (0.08)	**0.22 (0.02)**	
	PT	-0.12 (0.12)	-0.10 (0.11)	**0.18 (0.02)**	
	DU	0.00 (0.12)	-0.05 (0.11)	**0.17 (0.02)**	
	LW	-0.67 (0.17)	-0.65 (0.15)	**0.12 (0.03)**	
	LD	0.05 (0.26)	-0.02 (0.21)	**0.13 (0.04)**	
ABN	HY	-0.81 (0.03)	-0.77 (0.04)	0.15 (0.09)	**0.19 (0.02)**
	PT	-0.60 (0.08)	-0.52 (0.08)	-0.02 (0.12)	**0.17 (0.03)**
	DU	-0.78 (0.06)	-0.83 (0.04)	0.17 (0.12)	**0.24 (0.04)**
	LW	-0.84 (0.06)	-0.84 (0.05)	0.61 (0.17)	**0.25 (0.06)**
	LD	-0.25 (0.26)	-0.27 (0.21)	0.17 (0.22)	**0.15 (0.05)**

[1]MOT = sperm motility; PROMOT = sperm progressive motility; lnN_{cells} = log-transformed number of sperm cells per ejaculate; ABN = total morphological abnormalities.

[2]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

Within-Boar Variation Analysis

Variance component estimates from within-boar variation analysis are provided in Supplemental File S2. Table 4 shows heritability estimates for AVG, SD, and CV. For all

traits, except ABN in LW, the AVG showed the greatest heritability estimates compared with SD and CV. For MOT and ABN, the SD heritabilities were greater than CV heritabilities for all lines. For PROMOT, HY showed similar heritabilities for SD and CV. For PT and LD, SD heritabilities were lower than CV heritabilities. In contrast, DU and LW showed greater heritabilities for SD compared with CV. For lnN_{cells}, HY showed similar heritabilities for SD and CV. The PT, DU, and LW showed slightly greater heritabilities for SD compared with CV. The LD showed slightly lower heritability for SD compared with CV. The CV heritabilities for ABN in LW and LD (Table 4) were not considered different from 0 due to very small estimates of additive genetic variance compared with the residual variance (Supplemental File S2).

Negative genetic correlations between AVG and SD were found for the traits MOT and PROMOT (Table 5). The genetic correlations between AVG and SD were small and negligible for lnN_{cells}, whereas for ABN, they were moderately to highly positive (Table 5). In general, the SE were greater for the lines with fewer animals (LW and LD; Table 5), leading to greater uncertainty in the estimation of genetic correlations for these lines.

Table S2. Variance components, heritabilities (SE) from single-trait and variation analysis

Semen trait[1], population[2] and analysis type[3]	Number of animals[4]	Variance components[5]			Heritability
		σ_a^2	σ_{perm}^2	σ_ε^2	
MOT					
HY	2,708				
Single-trait analysis		8.24 (1.04)	8.59(0.71)	30.31 (0.09)	0.17 (0.02)

AVG		8.42 (1.08)		10.08 (0.75)	
SD		1.97 (0.35)		5.96 (0.30)	
CV		3.09 (0.65)		12.65 (0.60)	
PT	2,277				
Single-trait analysis		7.06 (1.11)	10.21 (0.85)	26.95 (0.09)	0.16 (0.02)
AVG		7.09 (1.14)		11.85 (0.89)	
SD		1.51 (0.34)		5.74 (0.32)	
CV		2.32 (0.61)		11.53 (0.59)	
DU	1,511				
Single-trait analysis		5.37 (1.07)	7.93 (0.79)	27.99 (0.12)	0.13 (0.02)
AVG		5.71 (1.14)		9.54 (0.85)	
SD		1.43 (0.34)		4.39 (0.30)	
CV		2.15 (0.58)		8.01 (0.52)	
LW	760				
Single-trait analysis		7.35 (2.16)	13.98 (1.77)	25.60 (0.18)	0.16 (0.04)
AVG		6.95 (2.10)		15.68 (1.78)	
SD		1.66 (0.58)		5.10 (0.52)	
CV		2.63 (1.17)		11.73 (1.12)	
LD	778				
Single-trait analysis		3.76 (1.14)	9.85 (1.05)	21.52 (0.16)	0.11 (0.03)
AVG		4.10 (1.15)		10.48 (1.04)	
SD		1.10 (0.34)		3.73 (0.33)	
CV		1.81 (0.62)		7.05 (0.62)	
PROMOT HY	2,708				

Single-trait					
analysis		14.33 (1.54)	10.44 (1.00)	40.78 (0.12)	0.22 (0.02)
AVG		14.63 (1.59)		12.09 (1.04)	
SD		1.46 (0.29)		6.09 (0.27)	
CV		3.73 (0.80)		15.45 (0.73)	
PT	2,277				
Single-trait					
analysis		13.03 (1.83)	14.60 (1.35)	41.46 (0.13)	0.19 (0.02)
AVG		12.71 (1.85)		17.11 (1.39)	
SD		1.08 (0.29)		5.55 (0.28)	
CV		3.20 (0.82)		13.15 (0.75)	
DU	1,511				
Single-trait					
analysis		10.50 (1.82)	11.06 (1.27)	40.42 (0.17)	0.17 (0.03)
AVG		10.87 (1.87)		12.93 (1.31)	
SD		0.82 (0.27)		4.48 (0.26)	
CV		2.20 (0.78)		12.24 (0.75)	
LW	760				
Single-trait					
analysis		13.93 (3.34)	17.91 (2.55)	37.53 (0.26)	0.20 (0.04)
AVG		13.58 (3.29)		20.02 (2.55)	
SD		1.24 (0.51)		4.98 (0.48)	
CV		2.96 (1.38)		14.80 (1.37)	
LD	778				
Single-trait					
analysis		7.59 (1.92)	13.47 (1.63)	35.26 (0.26)	0.13 (0.03)
AVG		7.54 (1.86)		14.77 (1.58)	
SD		0.08 (0.25)		4.51 (0.33)	
CV		0.77 (0.71)		10.84 (0.85)	

lnN_{cells}					
HY	2,708				
Single-trait analysis		0.03 (3.60E-03)	0.03 (2.50E-03)	0.08 (3.00E-04)	0.22 (0.02)
AVG		0.03 (3.60E-03)		0.04 (2.50E-03)	
SD		9.00E-04 (2.00E-04)		0.01 (2.00E-04)	
CV		0.08 (0.02)		0.49 (0.02)	
PT	2,277				
Single-trait analysis		0.03 (4.10E-03)	0.04 (3.10E-03)	0.09 (3.00E-04)	0.17 (0.02)
AVG		0.03 (3.90E-03)		0.04 (3.10E-03)	
SD		1.10E-03 (3.00E-04)		0.01 (3.00E-04)	
CV		0.09 (0.03)		0.51 (0.03)	
DU	1,511				
Single-trait analysis		0.02 (3.90E-03)	0.03 (2.80E-03)	0.10 (4.00E-04)	0.15 (0.02)
AVG		0.02 (3.80E-03)		0.03 (2.70E-03)	
SD		3.50E-03 (7.00E-04)		0.01 (5.00E-04)	
CV		0.30 (0.06)		0.58 (0.05)	
LW	760				
Single-trait analysis		0.01 (4.10E-03)	0.04 (3.80E-03)	0.07 (5.00E-04)	0.10 (0.03)
AVG		0.01 (3.80E-03)		0.04 (3.60E-03)	

		6.00E-04		0.01 (4.00E-	
SD		(4.00E-04)		04)	
CV		0.04 (0.03)		0.47 (0.04)	
LD	778				
Single-trait		0.02	0.03	0.07 (5.00E-	
analysis		(4.80E-03)	(4.10E-03)	04)	0.14 (0.04)
		0.02		0.04 (3.90E-	
AVG		(4.70E-03)		03)	
		3.00E-04		4.40E-03	
SD		(2.00E-04)		(3.00E-04)	
CV		0.03 (0.02)		0.38 (0.03)	
ABN					
HY	2,867				
Single-trait					
analysis		26.76 (3.55)	33.12 (2.54)	118.20 (0.53)	0.15 (0.02)
AVG		27.69 (3.64)		38.92 (2.61)	
SD		6.55 (1.10)		18.23 (0.91)	
		5077.40		287760.00	
CV		(5846.90)		(9286.70)	
PT	2,195				
Single-trait					
analysis		27.39 (4.57)	42.68 (3.61)	109.84 (0.59)	0.15 (0.02)
AVG		28.43 (4.58)		47.28 (3.61)	
SD		4.83 (1.14)		20.04 (1.08)	
		1227.50		65514.00	
CV		(1070.10)		(2215.40)	
DU	1,627				
Single-trait					
analysis		17.99 (3.99)	42.28 (3.34)	103.45 (0.68)	0.11 (0.02)
AVG		18.06 (3.96)		47.81 (3.31)	

SD		3.26 (0.91)		14.35 (0.86)	
		210.34		8230.70	
CV		(251.21)		(367.06)	
LW	879				
Single-trait analysis		17.79 (5.60)	47.16 (5.24)	139.97 (1.28)	0.09 (0.03)
AVG		19.23 (5.34)		50.62 (4.83)	
SD		7.81 (1.96)		16.88 (1.70)	
				3334500.00	
CV		0.83 (0.04)		(159150.00)	
LD	704				
Single-trait analysis		13.17 (6.19)	50.58 (6.18)	105.92 (1.16)	0.08 (0.04)
AVG		13.01 (5.74)		53.31 (5.65)	
SD		4.39 (1.90)		19.03 (1.92)	
		0.03		224720.00	
CV		(1.70E-03)		(11986.00)	

[1]MOT = sperm motility; PROMOT = sperm progressive motility; lnN_{cells} = log-transformed number of cells per ejaculate; ABN = total morphological abnormalities.

[2]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

[3]Single-trait analysis using repeatability animal model for MOT, PROMOT, lnN_{cells} and ABN or within-boar variation analysis (average [AVG], SD, and CV).

[4]Animals with ≥10 ejaculates.

[5]Variance components: σ_a^2 = additive genetic variance; σ_{perm}^2 = permanent environmental variance; σ_e^2 = residual variance.

Table 4. Heritabilities (SE) of variation traits for all semen traits and pig lines in the within-boar variation analysis

		Semen trait[3]			
Variation trait[1]	Pig line[2]	MOT	PROMOT	lnN_{cells}	ABN
AVG	HY	0.46 (0.05)	0.55 (0.05)	0.47 (0.04)	0.42 (0.05)
	PT	0.37 (0.05)	0.43 (0.05)	0.38 (0.05)	0.38 (0.05)
	DU	0.37 (0.07)	0.46 (0.07)	0.44 (0.06)	0.27 (0.06)
	LW	0.31 (0.08)	0.40 (0.09)	0.23 (0.07)	0.28 (0.07)
	LD	0.28 (0.07)	0.34 (0.07)	0.34 (0.08)	0.20 (0.08)
SD	HY	0.25 (0.04)	0.19 (0.04)	0.14 (0.04)	0.26 (0.04)
	PT	0.21 (0.04)	0.16 (0.04)	0.17 (0.04)	0.19 (0.04)
	DU	0.25 (0.05)	0.16 (0.05)	0.35 (0.06)	0.19 (0.05)
	LW	0.25 (0.08)	0.20 (0.08)	0.10 (0.06)	0.32 (0.07)
	LD	0.23 (0.07)	0.02 (0.05)	0.06 (0.05)	0.19 (0.08)
CV	HY	0.20 (0.04)	0.19 (0.04)	0.14 (0.04)	0.02 (0.02)
	PT	0.17 (0.04)	0.20 (0.05)	0.15 (0.04)	0.02 (0.02)
	DU	0.21 (0.05)	0.15 (0.05)	0.34 (0.06)	0.02 (0.03)
	LW	0.18 (0.08)	0.17 (0.08)	0.09 (0.06)	0.00 (0.00)
	LD	0.20 (0.07)	0.07 (0.06)	0.07 (0.05)	0.00 (0.00)

[1]AVG = average.

[2]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

[3]MOT = sperm motility; PROMOT = sperm progressive motility; lnN_{cells} = log-transformed number of sperm cells per ejaculate; ABN = total morphological abnormalities.

Table 5. Genetic correlations (SE) between average and SD for all lines and semen traits in the within-boar variation analysis

	Semen trait[2]			
Pig line[1]	MOT	PROMOT	lnN_{cells}	ABN
HY	-0.74 (0.05)	-0.83 (0.05)	-0.21 (0.12)	0.81 (0.04)
PT	-0.72 (0.07)	-0.80 (0.07)	0.08 (0.15)	0.95 (0.04)
DU	-0.54 (0.11)	-0.67 (0.11)	-0.18 (0.12)	0.78 (0.08)

LW	-0.55 (0.15)	-0.65 (0.13)	0.35 (0.37)	0.87 (0.07)
LD	-0.45 (0.16)	-0.99 (1.26)	-0.20 (0.34)	0.50 (0.21)

[1]HY = a synthetic line created around 1975 as a combination of the Large White and Pietrain pig breeds; PT = a Pietrain-based line; DU = a Duroc-based line; LW = a Large White-based line; LD = a Landrace-based line.

[2]MOT = sperm motility; PROMOT = sperm progressive motility; lnN_{cells} = log-transformed number of sperm cells per ejaculate; ABN = total morphological abnormalities.

DISCUSSION

The interest in studying the genetic control of semen traits has increased in recent years. With the advancement of molecular techniques, researchers have identified genes and genetic markers associated with semen characteristics (Xing et al., 2009; Gunawan et al., 2011, 2012; Kaewmala et al., 2011; Diniz et al., 2014). However, in many countries, young boars are selected for AI based on their breeding values regarding production traits and selecting boars based on semen traits is still not a common practice. Schulze et al. (2014) showed that 47.3% of young boars genetically proven for production traits were excluded from AI stations because some semen traits did not reach quality thresholds. In the Netherlands in 2014, this was only 15% of the young boars (AIM Varkens KI Nederland, M. L. W. J. Broekhuijse, personal communication), although this is still a significant number of animals. Commercially important production traits should not be the only criteria used for selection (Robinson and Buhr, 2005), because this may result in reduced semen production due to unfavorable genetic correlations between some production and semen traits (Oh et al., 2006a).

Selecting for semen traits is important for performance at AI stations whose objective is to maximize the number of insemination doses produced by each ejaculate. Therefore, the estimation of genetic parameters and the quantification of genetic variation for semen quality

and quantity allow an economic analysis of whether these traits should be included in breeding goal.

In this study, a large data set including 5 pig lines and 4 semen quality and quantity traits was used to estimate genetic parameters and the possibility of effective selection of boars based on these traits was shown. In addition, given that the AI industry prefers animals with low variation in semen traits between ejaculates, knowledge of heritabilities of within-boar variation of these traits is useful.

Heritabilities of moderate magnitude were estimated for all semen traits using multiple-trait analysis, indicating that selection for these traits could make reasonable progress. The 5 lines showed similar values for genetic parameters, except LD and also LW for some traits. One reason for these differences may be the lower genetic and residual variances estimated for semen traits in LD and also LW (for trait lnN_{cells}) compared with the other lines (Supplemental File S1). Moreover, the number of boars and ejaculates for those lines was lower in comparison with those for HY, PT, and DU, which may also lead to differences in genetic parameter estimates.

Large positive genetic correlations between MOT and PROMOT demonstrated a favorable genetic relation between these traits. The large negative genetic correlations observed between ABN and MOT and between ABN and PROMOT demonstrated that selection for a lower percentage of morphological defects would result in greater values of MOT and PROMOT, which is one of the goals of the AI industry. Based on the near-zero correlations, it seems that lnN_{cells} is genetically independent from MOT, PROMOT, and ABN, except for LW. A possible explanation for the nonzero genetic correlations in LW is the presence of pleiotropic QTL with higher minor allele frequencies in LW. These QTL alleles should have positive effects on lnN_{cells}, negative effects on MOT and PROMOT, and

positive effects on ABN. Higher minor allele frequencies of these QTL in LW could also explain the greater amount of the genetic variance in this line compared with the other lines.

Semen quality and quantity traits are repeatedly measured during the reproductive life of boars. In the literature, these traits are commonly analyzed using repeatability animal models (Wolf, 2009a,b, 2010; Wolf and Smital, 2009a,b; Diniz et al., 2014) similar to that performed in the present study. For MOT, we found greater heritabilities (except for LD) compared with our previous study (Diniz et al., 2014), in which we performed a genome-wide association study for MOT. In that study, a subset of the current data for LW and LD was analyzed, and the heritabilities were 0.11 for the LD and 0.14 for the LW (Table 6). For PROMOT, lnN_{cells}, and ABN, the average heritability across the 5 lines (0.22 ± 0.08, 0.16 ± 0.04, and 0.20 ± 0.04, respectively) was greater than the average reported in the literature when repeatability animal models were used (0.10 ± 0.04, 0.13 ± 0.04, and 0.11 ± 0.05, respectively; Table 6). The average genetic correlation estimated between PROMOT and ABN across lines (−0.65 ± 0.25) was stronger than estimates by Smital et al. (2005) and Wolf (2009b) for sire breeds and by Wolf (2009a), Wolf and Smital (2009b), and Wolf (2010) for

Czech Landrace (−0.34 ± 0.052, −0.59 ± 0.050, −0.51 ± 0.064, −0.57, and −0.48 ± 0.115, respectively) and weaker than estimates by Wolf (2009b) for Czech Large White, Czech Landrace, and dam breeds and by Wolf and Smital (2009a) and Wolf (2010) for Czech Large White (−0.87 ± 0.105, −0.92 ± 0.111, −0.93 ± 0.088, −0.88 ± 0.109, and −0.87 ± 0.076, respectively).

Comparing the number of boars and ejaculates used in our study with the data from these other studies, our data set is much larger, especially for HY and PT. This lead to more reliable genetic parameter estimates, evidenced by the lower SE. Other potential reasons for the dissimilarity between the studies are the different effects included in the statistical models,

the different breeds and crosses, and the different methods to measure semen traits, especially for progressive motility (CASA system or technician evaluation in a microscope).

Other approaches used to analyze semen data are the multiple-trait models (considering semen traits recorded at different ages as different traits) and random regression models (Oh et al., 2006b; Oh and See, 2008; Strathe et al., 2013). The random regression models provide a method for analyzing variance components that change over time (Oh and See, 2008), allowing modeling for both genetic and permanent effects as functions of time (Wolf, 2009b; Wolf and Smital, 2009a). In the present study, the average heritability for lnN_{cells} across lines (0.16 ± 0.04) was lower than those reported by Oh et al. (2006b) and Strathe et al. (2013), who applied random regression models to describe N_{cells} over the reproductive lifetime of boars (Table 6). Our finding was also lower than that of Oh and See (2008), who compared the random regression and multiple-trait animal models to study N_{cells} over the active lifetime of AI boars. However, care must be taken when interpreting results at the extremes of the period (Carabaño et al., 2007), and later performance may be harder to accurately predict from records collected at an early age (Oh et al., 2006b). According to Wolf (2009b) and Wolf and Smital (2009a), the classical animal models (as used in our study) require a lower number of parameters compared with random regression models.

In the present study, the estimation of genetic parameters for semen traits in the multiple-trait analysis (4×4) was performed using records from individual boar ejaculates. In contrast, in the within-boar variation analysis (when calculating AVG) and in Smital et al. (2005), genetic parameters were estimated based on the average values of individual boars rather than applying a repeatability model. When genetic parameters are estimated based on average values, heritability estimates increase because the residual variance estimates, which are divided by the harmonic mean of the number of ejaculates from each boar (Searle, 1971), are

lower. In addition, permanent environmental variance is not estimated separately in AVG analysis, which also inflates the AVG heritabilities. Therefore, as observed, additive genetic variances from the single-trait and AVG analyses were highly similar but residual variances from the AVG analysis were much lower (Supplemental File S2). This explains the lower heritability estimates from the repeatability model and similar heritability estimates from the AVG analysis compared with the results of Smital et al. (2005) for PROMOT, ABN and lnN_{cells} (Table 3, 4, and 6).

The average genetic correlation between PROMOT and ABN across lines was stronger (-0.65 ± 0.25) compared with that reported by Smital et al. (2005), who estimated a genetic correlation of -0.34 ± 0.052 between these traits. Low and negligible genetic correlations were found between PROMOT and lnN_{cells} and between ABN and lnN_{cells}, in agreement with the results of Smital et al. (2005), who estimated genetic correlations of -0.06 ± 0.058 and 0.14 ± 0.057, respectively, between these traits.

In the within-boar variation analysis, the AVG showed the greatest heritabilities followed by SD and CV for some traits. For selection purposes, a trait with greater heritability is preferred; therefore, in the variation study, the AVG and SD in a multiple-trait analysis (2 × 2) could be more suitable than CV. Moreover, selection for low CV may include animals with low AVG if their SD is sufficiently low as well, which is not favorable. In this study, a multiple-trait analysis (3 × 3) of AVG, SD, and CV was not performed because of the functional dependency between CV and SD and between CV and AVG (CV = (SD/AVG) × 100). According to Wolf (2009a) and Wolf (2010), traits that have functional relationships should not be analyzed together in a multiple-trait model because of the full dependency between the evaluated traits. In addition, one of our aims was to access the ability of variation traits to quantify uniformity by comparing the SD and CV separately.

A negative genetic correlation between AVG and SD was found for the traits MOT and PROMOT, validating prior interest in animals with high AVG and low SD for these traits. The same favorable result was found for ABN, with positive genetic correlations between AVG and SD, because for ABN, the interest is in animals with low AVG and also low SD. For $lnN_{cells,}$ the genetic correlations between AVG and SD were negligible. The values of the traits MOT and PROMOT are generated by the CASA system, which analyses between 300 and 1,000 sperm cells of a single ejaculate. This finite sample of cells may affect the phenotypic correlation of the AVG and SD of ejaculates within a boar. As a result, the genetic correlations between AVG and SD may be influenced in a negative or positive direction. However, we assume the "true" AVG and SD of ejaculates from a single boar to be determined by his breeding values for AVG and SD, which are assumed normally distributed across animals. Therefore, these breeding values for the AVG and SD of the same trait may have a positive, negative, or no genetic correlation, as we estimated in this study.

In the literature, the estimation of genetic parameters for within-boar variation of semen quality and quantity traits has not yet been reported. The majority of studies describing the estimation of genetic parameters for mean and variation of other traits considered the AVG, SD, and/or CV but in single-trait analyses. For example, CV, SD, and other variation traits were used to study the within-litter variation of piglet weights (Hermesch et al., 2001; Damgaard et al., 2003; Wolf et al., 2008) and the variation of somatic cell score in dairy cattle (Urioste et al., 2010; Wijga et al., 2012). The authors also estimated the genetic parameters for harmonic and arithmetic means of piglet weights and for lactation-average somatic cell score using a single-trait approach. In all studies, the authors found greater heritabilities for mean traits than for variation traits, which is in accordance with the greater heritabilities estimated for the AVG compared with SD and CV in the present study.

Moderate values of heritability indicate the possibility of effective boar selection based on semen quality and quantity traits. Relevant genetic variation exists for these traits, both in level and in within-boar variation. This quantification allows an economic analysis of whether these traits should be included in breeding goals. Given the estimated genetic correlations and heritabilities for within-boar variation traits, the AVG and SE (in a 2 × 2 multiple-trait analysis) are the best traits to consider when selecting for reduced variation in semen characteristics. In future studies, determining the effect of semen traits on female fertility traits (such as the total number of piglets born and the number born alive), as well as the genetic correlation among these traits, is of great interest and could aid in the implementation of semen characteristics in boar breeding programs.

Table 6. Heritability estimates for boar semen quality and quantity traits from different studies

Semen trait[1]	Number of boars	Number of ejaculates	Heritability (SE)	Notes	Authors
MOT	760	32,884	0.11	Landrace-based dam line	Diniz et al. (2014)
	645	32,576	0.14	Large-White based dam line	
PROMOT	2,862	210,733	0.38 (0.04)	Mean value of SE according to different animal models. Genetic parameters estimated based on average values of individual boars	Smital et al. (2005)
	2,077	151,755	0.05 (0.01)	Repeatability animal model	Wolf (2009a)
				Repeatability animal model	
	615	26,017	0.06 (0.02)	Czech Large White	Wolf (2009b)
	653	36,747	0.16 (0.03)	Czech Landrace	
	1,268	62,764	0.14 (0.02)	Dam breeds	
	2,407	153,066	0.08 (0.01)	Sire breeds	
	1,417	75,567	0.13 (0.02)	Repeatability animal model	Wolf and Smital (2009a)
	2,077	151,755	0.05	Repeatability animal model	Wolf and Smital (2009b)
	778	37,137	0.08 (0.02)	Repeatability animal model Czech Large White	Wolf (2010)

	841	51,341	0.12 (0.02)	Czech Landrace	
N_{cells}	2,862	210,733	0.42 (0.03)	Mean value of SE according to different animal models. Genetic parameters estimated based on average values of individual boars	Smital et al. (2005)
	834	19,629	0.27 to 0.48	Analysis were carried out using random regression models	Oh et al. (2006b)
	750		0.26 (0.03) to 0.79 (0.09)	Analysis carried out using multitrait approach, considering 9, 12,15,18,21, 24 and 27 months of age as separate traits	Oh and See (2008)
	2,077	151,755	0.17 (0.01)	Repeatability animal model	Wolf (2009a)
	1,417	75,567	0.10 (0.02)	Repeatability animal model	Wolf and Smital (2009a)
	2,077	151,755	0.17	Repeatability animal model	Wolf and Smital (2009b)
	778	37,137	0.10 (0.03)	Repeatability animal model Czech Large White	Wolf (2010)
	841	51,341	0.12 (0.02)	Czech Landrace	
	3,145	95,267	0.18 to 0.20	Analysis were carried out using random regression models	Strathe et al. (2013)
ABN	2,862	210,733	0.34 (0.03)	Mean value of SE according to different animal models. Genetic parameters estimated based on average values of individual boars	Smital et al. (2005)
	615	26,017	0.04 (0.02)	Repeatability animal model Czech Large White	
	653	36,747	0.12 (0.04)	Czech Landrace	Wolf (2009b)
	1,268	62,764	0.06 (0.01)	Dam breeds	
	2,407	153,066	0.17 (0.02)	Sire breeds	
	2,077	151,755	0.15 (0.02)	Repeatability animal model	Wolf (2009a)
	1,417	75,567	0.07 (0.03)	Repeatability animal model	Wolf and Smital (2009a)
	2,077	151,755	0.16	Repeatability animal model	Wolf and Smital (2009b)
	778	37,137	0.12 (0.02)	Repeatability animal model Czech Large White	Wolf (2010)
	841	51,341	0.10 (0.03)	Czech Landrace	

[1]MOT = sperm motility; PROMOT = sperm progressive motility; N_{cells} = number of sperm cells per ejaculate; ABN = total morphological abnormalities.

LITERATURE CITED

Broekhuijse, M. L. W. J., H. Feitsma, and B. M. Gadella. 2011a. Field data analysis of boar semen quality. Reprod. Domest. Anim. 46: 59-63. doi: 10.1111/j.1439-0531.2011.01861.x

Broekhuijse, M. L. W. J., E. Šoštarić, H. Feitsma, and B. M. Gadella. 2011b. Additional value of computer assisted semen analysis (CASA) compared to conventional motility assessments in pig artificial insemination. Theriogenology 76: 1473-1486. e1471. doi: 10.1016/j.theriogenology.2011.05.040

Carabaño, M. J., C. Díaz, C. Ugarte, and M. Serrano. 2007. Exploring the use of random regression models with legendre polynomials to analyze measures of volume of ejaculate in Holstein bulls. J. Dairy Sci. 90: 1044-1057. doi: 10.3168/jds.S0022-0302(07)71591-6

Damgaard, L. H., L. Rydhmer, P. Løvendahl, and K. Grandinson. 2003. Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. J. Anim. Sci. 81: 604-610. doi:10.2527/2003.813604x

Diniz, D. B., M. S. Lopes, M. L. W. J. Broekhuijse, P. S. Lopes, B. Harlizius, S. E. F. Guimarães, N. Duijvesteijn, E. F. Knol, and F. F. Silva. 2014. A genome-wide association study reveals a novel candidate gene for sperm motility in pigs. Anim. Reprod. Sci. 151: 201-207. doi: 10.1016/j.anireprosci.2014.10.014

Gilmour, A. R., B. J. Gogel, B. R. Cullis, R. Thompson, and D. Butler. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.

Gunawan, A., M. U. Cinar, M. J. Uddin, K. Kaewmala, D. Tesfaye, C. Phatsara, E. Tholen, C. Looft, and K. Schellander. 2012. Investigation on association and expression of

ESR2 as a candidate gene for boar sperm quality and fertility. Reprod. Domest Anim. 47: 782-790. doi: 10.1111/j.1439-0531.2011.01968.x

Gunawan, A., K. Kaewmala, M. J. Uddin, M. U. Cinar, D. Tesfaye, C. Phatsara, E. Tholen, C. Looft, and K. Schellander. 2011. Association study and expression analysis of porcine ESR1 as a candidate gene for boar fertility and sperm quality. Anim. Reprod. Sci. 128: 11-21. doi: 10.1016/j.anireprosci.2011.08.008

Hermesch, S., B. G. Luxford, and H. -U. Graser. 2001. Genetic parameters for piglet mortality, within litter variation of birth weight, litter size and litter birth weight. Proc. Assoc. Adv. Anim. Breed. Genet. 14:211-214.

Kaewmala, K., M. J. Uddin, M. U. Cinar, C. Große-Brinkhaus, E. Jonas, D. Tesfaye, C. Phatsara, E. Tholen, C. Looft, and K. Schellander. 2011. Association study and expression analysis of CD9 as candidate gene for boar sperm quality and fertility traits. Anim. Reprod. Sci. 125: 170-179. doi: 10.1016/j.anireprosci.2011.02.017

Oh, S. H., and M. T. See. 2008. Comparison of genetic parameter estimates of total sperm cells of boars between random regression and multiple trait animal models. Asian-Aust. J. Anim. Sci. 21: 923-927. doi: 10.5713/ajas.2008.70383

Oh, S. H., M. T. See, T. E. Long, and J. M. Galvin. 2006a. Estimates of genetic correlations between production and semen traits in boar. Asian-Aust. J. Anim. Sci. 19: 160-164. doi: 10.5713/ajas.2006.160

Oh, S. H., M. T. See, T. E. Long, and J. M. Galvin. 2006b. Genetic parameters for various random regression models to describe total sperm cells per ejaculate over the reproductive lifetime of boars. J. Anim. Sci. 84: 538-545. doi: 10.2527/2006.843538x

Robinson, J. A. B., and M. M. Buhr. 2005. Impact of genetic selection on management of boar replacement. Theriogenology 63: 668-678. doi: 10.1016/j.theriogenology.2004.09.040

Schulze, M., S. Buder, K. Rüdiger, M. Beyerbach, and D. Waberski. 2014. Influences on semen traits used for selection of young AI boars. Anim. Reprod. Sci. 148: 164-170. doi: 10.1016/j.anireprosci.2014.06.008

Searle, S.R. 1971. Linear Models. Wiley, New York.

Smital, J., J. Wolf, and L. L. De Sousa. 2005. Estimation of genetic parameters of semen characteristics and reproductive traits in AI boars. Anim. Reprod. Sci. 86: 119-130. doi: 10.1016/j.anireprosci.2004.05.023

Strathe, A. B., I. H. Velander, T. Mark, T. Ostersen, C. Hansen, and H. N. Kadarmideen. 2013. Genetic parameters for male fertility and its relationship to skatole and androstenone in Danish Landrace boars. J. Anim. Sci. 91: 4659-4668. doi: 10.2527/jas.2013-6454

Urioste, J. I., J. Franzén, and E. Strandberg. 2010. Phenotypic and genetic characterization of novel somatic cell count traits from weekly or monthly observations. J. Dairy Sci. 93: 5930-5941. doi: 10.3168/jds.2010-3301

Wijga, S., J. W. M. Bastiaansen, E. Wall, E. Strandberg, Y. de Haas, L. Giblin, H. Bovenhuis. 2012. Genomic associations with somatic cell score in first-lactation Holstein cows. J. Dairy Sci. 95: 899-908. doi: 10.3168/jds.2011-4717

Wolf, J. 2009a. Genetic correlations between production and semen traits in pig. Animal 3: 1094-1099. doi: 10.1017/S1751731109004686

Wolf, J. 2009b. Genetic parameters for semen traits in AI boars estimated from data on individual ejaculates. Reprod. Domest. Anim. 44: 338-344. doi: 10.1111/j.1439-0531.2008.01083.x

Wolf, J. 2010. Heritabilities and genetic correlations for litter size and semen traits in Czech Large White and Landrace pigs. J. Anim. Sci. 88: 2893-2903. doi: 10.2527/jas.2009-2555

Wolf, J., and J. Smital. 2009a. Effects in genetic evaluation for semen traits in Czech Large White and Czech Landrace boars. Czech J. Anim. Sci. 54: 349-358.

Wolf, J., and J. Smital. 2009b. Quantification of factors affecting semen traits in artificial insemination boars from animal model analyses. J. Anim. Sci. 87: 1620-1627. doi: 10.2527/jas.2008-1373

Wolf, J., E. Žáková, and E. Groeneveld. 2008. Within-litter variation of birth weight in hyperprolific Czech Large White sows and its relation to litter size traits, stillborn piglets and losses until weaning. Livest. Sci. 115: 195-205. doi: 10.1016/j.livsci.2007.07.009

Xing, Y., J. Ren, D. Ren, Y. Guo, Y. Wu, G. Yang, H. Mao, B. Brenig, L. Huang. 2009. A whole genome scanning for quantitative trait loci on traits related to sperm quality and ejaculation in pigs. Anim. Reprod. Sci. 114: 210-218. doi: 10.1016/j.anireprosci.2008.08.008

CHAPTER 3

Weighted single-step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs

Daniele B. D. Marques[1], John W. M. Bastiaansen[2], Marleen L. W. J. Broekhuijse[3], Marcos S. Lopes[3,4], Egbert F. Knol[3], Barbara Harlizius[3], Simone E. F. Guimarães[1], Fabyano F. Silva[1], Paulo S. Lopes[1]

[1] Universidade Federal de Viçosa, Animal Science Department, 36.570-000, Viçosa-MG, Brazil, [2] Wageningen University & Research, Animal Breeding and Genomics, P.O. Box 338 6700 AH, Wageningen, the Netherlands; [3] Topigs Norsvin Research Center B.V., P.O. Box 43, 6640 AA Beuningen, the Netherlands; [4] Topigs Norsvin, 80.420-210, Curitiba-PR, Brazil.

Genetics Selection Evolution (submitted)

Abstract

Background: In recent years, the interest in studying the molecular processes affecting semen traits has increased. Our aim with this study was to identify quantitative trait loci (QTL) regions associated with four semen traits (motility, progressive motility, number of sperm cells per ejaculate and total morphological defects) in two commercial pig lines (L1: Large White type and L2: Landrace type). A weighted single-step genome-wide association study (WssGWAS) was conducted as the number of animals with both phenotypes and genotypes was low in our data set (L1=349, L2=446) and because it allows the updates of different weights for SNPs, which are used to construct G matrices, resulting in improved precision in estimating GEBVs and SNP effects. Additionally, we also aimed to identify candidate genes within QTL regions that explained the highest proportions of genetic variance. Subsequently, we performed gene network analyses to investigate the biological processes shared by genes identified for the semen traits in the two lines.

Results: QTL regions explaining up to 10.8% of the genetic variance of the semen traits were identified in 12 chromosomes in L1 and in 11 chromosomes in L2. A total of 16 QTL regions in L1 and 6 QTL regions in L2 were associated with two or more traits within population. Candidate genes *SCN8A*, *PTGS2*, *PLA2G4A*, *DNAI2*, *IQCG* and LOC102167830 were identified in L1 and candidate genes *NME5*, *AZIN2*, *SPATA7*, *METTL3* and *HPGDS* were identified in L2. No regions were found to overlap between the lines. However, the gene network analysis for progressive motility revealed two genes in L1 (*PLA2G4A* and *PTGS2*) and one gene in L2 (*HPGDS*) sharing the biological processes of eicosanoid biosynthesis and arachidonic acid metabolism. *PTGS2* and *HPGDS* also shared the biological process of cyclooxygenase pathway.

Conclusions: We identified several QTL regions associated with semen traits in two pig lines, confirming the assumption of genetic complexity of these traits. A large part of the genetic variance of the semen traits under study is explained by different genes in the two evaluated lines. Nevertheless, the gene network analysis showed candidate genes sharing biological pathways that occur in mammalian testes, in both the lines.

Keywords: pig; single-step; GWAS; semen quality; gene network

Background

Artificial insemination (AI) pig industry mainly focuses on maximizing the number of insemination doses produced by each boar ejaculate. To achieve this goal, the ability of boars to produce high-quality semen (high motility and progressive motility with low level of morphological defects) in sufficient quantity (high number of sperm cells per ejaculate) is decisive [1].

In recent years, with the fast advances in high-throughput genotyping and in molecular techniques in general, the interest in studying the molecular processes and genetic mechanisms affecting semen traits has increased. Genes and markers associated with pig semen traits have been described in literature [2-8]. However, there are hardly any studies analyzing large datasets to identify novel quantitative trait loci (QTL) and to provide a deeper knowledge of genes controlling boar semen traits. One reason for this situation is the genetic complexity of the process of sperm cells production and maturation. Mammalian spermatogenesis requires coordination among different genes and cell types (germ cells, Sertoli cells, Leydig cells) [9]. The process occurs in the seminiferous tubules of testes in three steps: mitotic phase, meiotic phase, and spermiogenesis [10]. In the first step (mitosis),

spermatogonias produce primary spermatocytes, which enter the first stage of meiosis (meiosis I), producing secondary spermatocytes. Then haploid round spermatids are generated after the second step of meiosis (meiosis II). In the last phase, spermiogenesis, the spermatids undergoes morphological transformations, getting spermatozoa shape. Then, the new pre-formed spermatozoa go through epididymis to maturate and acquire motility [10]. Mutations and impaired expression of genes controlling the whole process of spermatogenesis and sperm maturation can lead to several semen quality and fertility problems.

Genome-wide association studies (GWAS) are commonly used to identify single nucleotide polymorphisms (SNPs) associated with QTL with major effect [11]. The weighted single-step GWAS (WssGWAS), proposed by Wang et al. [12], is a iterative method that allows estimation of SNP effects using genomic estimated breeding values (GEBVs) from single-step GBLUP (ssGBLUP, [13]) on all phenotyped, genotyped and pedigree-related animals. In addition, it allows the updates of different weights for SNPs at each iteration, which are used to construct G matrices, resulting in improved precision in estimating GEBVs and SNP effects [12]. Therefore, when the number of animals with both phenotypes and genotypes is low and the traits are controlled by few QTL with large effects, the WssGWAS may perform better. Recent studies have used this method for production traits in livestock [14-20], but not for boar semen traits.

In a post-GWAS study, a gene network analyses can be performed for candidate genes related to QTL regions identified in GWAS. The gene network is used to investigate the shared pathways and biological processes involving these genes [21]. In addition, the biological information provided by these gene networks are helpful to understand the genetic differences between populations for the same trait [22].

Our aim with this study was to identify QTL regions associated with four semen traits (motility, progressive motility, number of sperm cells per ejaculate and total morphological defects) in pigs. Additionally, we also aimed to identify candidate genes within those QTL regions that explained the highest proportions of genetic variance. To achieve our goal, we performed a weighted single-step GWAS in two commercial pig lines (L1: Large White type and L2: Landrace type). Subsequently, we performed gene network analyses to investigate the biological processes shared by genes identified for the semen traits in the two lines evaluated in this study.

Methods

Phenotypic, genotypic and pedigree data

Phenotypic data were available from two commercial pig lines: a Large White type line (L1) and a Landrace type line (L2). The traits evaluated were: 1) sperm motility (MOT), which is the proportion of moving sperm cells in an ejaculate; 2) sperm progressive motility (PROMOT), defined as the proportion of sperm cells moving in a straight line; 3) abnormal sperm cells (ABN), which are total number of sperm cells with morphological abnormalities; and 4) total number of sperm cells in the ejaculate (Ncells in 10^6 sperm cells). All semen traits were assessed on the day of semen collection using fresh semen.

The ejaculate samples were collected between January 2007 and October 2014. MOT and PROMOT were evaluated using the UltiMateTM CASA system (Hamilton Thorne Inc., Beverly, MA, USA). Ncells was calculated as the product of the semen volume (mL) and concentration (10^6 mL^{-1}, measured by the CASA system). It was not normally distributed, and was log-transformed before further analyses (lnNcells). Ejaculates evaluated for ABN were analyzed microscopically at a magnification x 1000 by a trained technician in a phase

contrast microscope and a thermal plate (BH-2, Olympus, Tokyo, Japan), counting 100 sperm cells per assessment.

The phenotypic data for MOT, PROMOT and lnNcells consisted of 43,455 ejaculates for L1 (866 boars) and 39,161 ejaculates for L2 (900 boars). For ABN, the phenotypic data consisted of 13,366 ejaculates for L1 (849 boars) and 9,853 ejaculates for L2 (886 boars). The average number of ejaculates per boar for MOT, PROMOT and lnNcells were 50.18 ± 38.12 for L1 and 43.51 ± 36.37 for L2. For ABN, the average number of ejaculates per boar were 15.74 ± 11.19 for L1 and 11.12 ± 8.99 for L2. Mean, standard deviation, minimum and maximum values of semen traits in L1 and L2 are shown in Table 1.

Genotypic data of 3,737 animals (856 males and 2,881 females) from L1 and 3,307 animals (953 males and 2,354 females) from L2 were available. A majority of the animals (2,718 for L1 and 2,394 for L2) were genotyped using the Illumina PorcineSNP60 BeadChip (Illumina, Inc., San Diego, CA). Part of animals from L1 (n=1,019) and L2 (n= 913) were genotyped using the (Illumina, Inc.) GeneSeek Custom 80K SNP chip (GeneSeek Inc., Lincoln, NE). Quality control was performed excluding SNPs that had unknown position on the Pig genome build10.2 [23], were located on sex chromosomes, had call rate ≤ 0.95, had minor allele frequency <0.01 or had strong deviation from Hardy-Weinberg equilibrium ($\chi 2$ values > 600). Animals with frequency of missing genotypes ≥ 0.05 were also excluded. After quality control, the software Beagle version 3.3.2 [24] was used to input missing genotypes of the animals genotyped with the SNP60 BeadChip. In addition, genotypes from the animals genotyped with the GeneSeek Custom 80K SNP chip were imputed to the set of SNP on the SNP60 BeadChip that passed the quality control. The

total number of SNPs remaining after quality control and used for GWAS analysis were 39,945 SNPs in L1 and 41,478 SNPs in L2.

The complete pedigree included a total of 8,352 animals for L1 and 8,271 animals for L2. The total number of animals that remained in the pedigree file after pedigree pruning were 6,724 for L1 and 6,502 for L2. Most animals had either phenotypic or genotypic data. The number of animals with both phenotypes and genotypes was 349 for L1 and 446 for L2.

Table 1 Descriptive statistics of semen traits

	Mean	SD	Min	Max
MOT				
L1	86.5	7.1	10	100
L2	87.1	6.5	14	100
PROMOT				
L1	78.6	8.3	0	100
L2	77.4	7.9	0	100
lnN_{cells}				
L1	25.0	0.4	24.0	26.4
L2	24.9	0.4	22.5	26.4
ABN				
L1	19.3	14.8	1	98
L2	14.4	12.6	1	99

Mean, standard deviation (SD), minimum (min), maximum (max) values of semen traits in two pig lines (L1= Large White and L2= Landrace).

Semen traits: MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities.

Statistical analysis

The weighted ssGBLUP analysis was conducted using the BLUPF90 software family [25] adapted for genomic analyses [26]. Variance components were first estimated in AIREMLF90 and then used in BLUPF90 to predict genomic breeding values. SNP effects were then calculated using postGSf90 software.

The single-trait animal model for ssGBLUP was the following:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Za} + \boldsymbol{Wp} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{y}$ is the vector of phenotypic observations; $\boldsymbol{\beta}$ is the vector of fixed effects (combined effects of AI station, year and month of semen collection; the laboratory where the samples were analyzed and the covariates interval between two subsequent semen collections in days and age of the boar in months at the semen collection); $\boldsymbol{a}$ is the vector of random additive genetic effects; $\boldsymbol{p}$ is the vector of random permanent environmental effects; $\boldsymbol{\varepsilon}$ is the vector of random residuals; and **X**, **Z** and **W** are the incidence matrices of $\boldsymbol{\beta}$, $\boldsymbol{a}$ and $\boldsymbol{p}$, respectively.

It was assumed that $\boldsymbol{a} \sim \mathrm{N}(\mathbf{0}, \mathbf{H}\sigma_a^2)$, $\boldsymbol{p} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}\sigma_p^2)$ and $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}\sigma_e^2)$, where σ_a^2 , σ_p^2 and σ_e^2 are the additive genetic, permanent environmental and residual variances, respectively; **H** is the matrix which combines pedigree and genomic information [13] and **I** is an identity matrix. The inverse of **H** matrix needed for mixed model equations is given by:

$$\boldsymbol{H^{-1}} = \boldsymbol{A^{-1}} + \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{G^{-1}} - \boldsymbol{A_{22}^{-1}} \end{bmatrix},$$

where $\boldsymbol{A}$ is the numerator relationship matrix based on pedigree for all animals; $\boldsymbol{A_{22}}$ is the numerator relationship matrix for genotyped animals; and $\boldsymbol{G}$ is the genomic relationship matrix [27]. This matrix was obtained as follows:

$$\boldsymbol{G} = \frac{\boldsymbol{ZDZ'}}{\sum_{i=1}^{M} 2p_i(1-p_i)} \ ,$$

where $\boldsymbol{Z}$ is a matrix of gene content adjusted for allele frequencies (0, 1 or 2 for aa, Aa and AA, respectively) ; $\boldsymbol{D}$ is a diagonal matrix of weights for SNP variances (initially $\boldsymbol{D} = \boldsymbol{I}$); M is the number of SNPs and p_i is the minor allele frequency of ith SNP.

SNP effects and weights for WssGWAS were obtained according to Wang et al. [12] in following steps:

1. In the first step (t=1): $\boldsymbol{D} = \boldsymbol{I}$; $\boldsymbol{G}_{(t)} = \boldsymbol{ZD}_{(t)}\boldsymbol{Z'}\lambda$, where t is the iteration number and $\lambda = \frac{1}{\sum_{i=1}^{M} 2p_i(1-p_i)}$ [27];

2. GEBVs were calculated for entire data set using ssGBLUP;

3. GEBVs were converted to SNP effects ($\hat{u}$): $\hat{u}_{(t)} = \lambda \boldsymbol{D}_{(t)}\boldsymbol{Z'}\boldsymbol{G}_{(t)}^{-1}\,\hat{\boldsymbol{a}}_g$, where $\hat{\boldsymbol{a}}_g$ is the GEBVs of animals which were also genotyped;

4. Weight for each SNP was calculated as: $d_{i(t+1)} = \hat{u}_{i(t)}^2 2p_i(1\text{-}p_i)$, where i is the i-th SNP;

5. SNP weights were normalized to keep the total genetic variance constant:

$D_{(t+1)} = \frac{\mathrm{tr}(D_{(1)})}{\mathrm{tr}(D_{(t+1)})} D_{(t+1)}$;

6. $\boldsymbol{G}_{(t+1)} = \boldsymbol{ZD}_{(t+1)}\boldsymbol{Z'}\lambda$ was calculated;

7. $t = t + 1$;

8. Loop to step 2.

The ideal number of iterations was chosen based on the realized accuracies of GEBVs according to Legarra et al. [28] and performed by Wang et al. [14]. Based on that, three iterations were used, and in the first iteration the weights were not applied ($\boldsymbol{D} = \boldsymbol{I}$). Each iteration updates weights for SNPs, which were used to construct G matrices and improve GEBVs estimates and, consequently, SNP effects. The percentage of genetic variance

explained by *i*th set of consecutive SNPs (*i*th SNP window) was calculated as described by Wang et al. [14]:

$$\frac{Var\ (a_i)}{\sigma_a^2}\ x\ 100\% = \frac{Var\left(\sum_{j=1}^{x} Z_j \hat{u}_j\right)}{\sigma_a^2}\ x\ 100\%,$$

where a_i is the genetic value of *i*th SNP window that consists of a region of consecutive SNP located within 0.4 Mb, being 0.4 Mb the average haplotype block size in commercial pig lines [29], including the lines considered in the present study; σ_a^2 is the total genetic variance; Z_j is the vector of gene content of the *j*th SNP for all individuals and $\hat{u}_j$ is the effect of the *j*th SNP within the *i*th window. Manhattan plots showing these windows were created using R software.

Search for candidate genes and gene network analyses

The search for genes within relevant QTL regions identified from WssGWAS was performed in three steps. Firstly, for the four traits (MOT, PROMOT, lnNcells and ABN), the SNP windows that explained 1% or more of total genetic variance were selected within each line (L1 and L2). Then, a search for overlapping windows for two or more traits of the same line was performed. Windows were considered to overlap if their midpoints were less than 0.4 Mb apart. In a second step, the three most important windows (that explained the highest proportion of genetic variance) for each trait and the region 0.4 Mb apart from these windows midpoints were also identified. Thirdly, based on selected windows in the first step (>1% of variance explained), overlapping windows for the same traits but in the two different lines were investigated (also considering maximum distance between midpoints of 0.4 Mb). After this third step, based on the starting and ending position of each identified and selected

QTL region, genes in these QTL regions were identified in Gene database for *Sus scrofa* available at "National Center for Biotechnology Information" (NCBI – http://www.ncbi.nlm.nih.gov). All genes identified were investigated for their relationship with the traits under study in literature by a manual search. In addition, based on human genes with the same description, four gene network analyses describing the biological processes and relations between L1 and L2 set of genes identified for the same traits were performed using the ClueGO and CluePedia Cytoscape plug-ins [30, 31]. The ClueGO combines Gene Ontology (GO) terms and KEGG/BioCarta pathways, developing a GO/pathway network. It also calculates enrichment and depletion tests for groups of genes based on hypergeometric distribution and corrects the *P*-values for multiple testing [30]. Using the CluePedia, new associations between genes can be discovered with enrichments and added to ClueGO pathways [31].

Results

Low to moderate heritabilities ranging from 0.10 to 0.31 were found. Variance components are shown in Additional file 1: Table S1. The heritabilities were higher for MOT, PROMOT and ABN in L1 (Table 2). Searches for candidate genes were performed within relevant QTL regions, which were: 1) overlapping regions (in two or more traits) with proportion of variance explained ≥ 1% and 2) the three windows (and the region 0.4 Mb from the windows midpoints) that explained the highest proportion of genetic variance for each trait. For L1, relevant QTL regions were found on *Sus scrofa* chromosome 1 (SSC1), SSC3, SSC4, SSC5, SSC6, SSC8, SSC9, SSC10, SSC12, SSC13, SSC14 and SSC15 (Table 3). For L2, they were located on SSC1, SSC2, SSC3, SSC6, SSC7, SSC8, SSC9, SSC11, SSC13, SSC15 and SSC18 (Table 4). For L1, the three most important windows for MOT, PROMOT,

lnNcells and ABN explained 14.39%, 18.16%, 15.67% and 18.78% of the genetic variance, respectively (Fig. 1; Additional file 2: Table S2). For L2, the three most important windows explained 21.88%, 18.68%, 18.29% and 13.81% of the genetic variance of each trait, respectively (Fig. 2; Additional file 3: Table S3).

A total of 110 genes located in the relevant QTL regions were identified in L1. Among these, six genes related to MOT, PROMOT, lnNcells and ABN were described in the literature (Table 3). For L2, 106 genes were found in the detected relevant regions and five genes were related to MOT, PROMOT, lnNcells and ABN (Table 4).

Overlapping regions between L1 and L2 for the same traits were not found in this study. Nevertheless, the gene network analysis for PROMOT revealed two genes in L1 (*PLA2G4A* and *PTGS2*) and one gene in L2 (*HPGDS*) sharing the biological processes of eicosanoid biosynthesis and arachidonic acid metabolism. *PTGS2* in L1 and *HPGDS* in L2 were also found sharing the biological process of cyclooxygenase pathway (Fig. 3).

Table S1 Variance components and standard errors (in parenthesis) for semen traits

Trait[a]	Pig line	σ^2_a (SE)	σ^2_{perm} (SE)	σ^2_e (SE)
MOT	L1	14.02 (3.98)	27.45 (3.18)	26.16 (0.18)
	L2	5.31 (1.83)	17.89 (1.66)	21.68 (0.16)
PROMOT	L1	28.26 (5.74)	26.01 (3.88)	38.14 (0.26)
	L2	7.77 (2.30)	20.32 (2.00)	35.64 (0.26)
lnN_{cells}	L1	0.01 (0.0038)	0.04 (0.0036)	0.07 (0.0005)
	L2	0.02 (0.0049)	0.04 (0.0043)	0.07 (0.0005)
ABN	L1	57.40 (14.73)	91.12 (11.51)	110.38 (1.45)
	L2	25.45 (9.10)	75.76 (8.06)	75.47 (1.17)

[a] Semen traits. MOT = sperm motility; PROMOT = sperm progressive motility; lnNcells = number of sperm cells per ejaculate; ABN = total morphological abnormalities.

[b] L1 = Large White type line; L2 = Landrace type line

σ_a^2 = additive genetic variance; σ_{perm}^2 = permanent environmental variance; σ_e^2 = residual variance.

Table 2 Estimated heritabilities (standard error) of the evaluated semen traits

Semen traits[a]	Pig lines[b]	
	L1	L2
MOT	0.21 (0.05)	0.12 (0.04)
PROMOT	0.31 (0.05)	0.12 (0.03)
lnNcells	0.10 (0.03)	0.13 (0.03)
ABN	0.22 (0.05)	0.14 (0.05)

[a]MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities.

[b]L1: Large White type and L2: Landrace type.

Table 3 Individual and overlapping QTL regions for semen traits in L1

Chr[a]	QTL region (Mb)[b]	Var (%)[c] MOT[d]	Var (%) PROMOT	Var (%) lnNcells	Var (%) ABN	Candidate gene[f]
1	135.51-136.31	1.2	1.2	-[e]	6.0	-[g]
1	255.48-256.28	1.2	-	-	1.5	-
1	290.90-291.84	1.0	-	-	1.3	-
1	305.18-305.98	-	-	4.6	-	-
3	28.53-29.33	1.0	-	-	2.9	-
4	28.25-29.05	2.5	6.4	-	5.9	-
4	84.90-85.73	1.4	2.5	-	-	-
4	123.12-124.20	1.8	1.0	-	1.3	-
5	17.61-18.47	7.5	3.0	-	-	*SCN8A*
6	8.24-9.13	2.6	1.1	-	-	-
8	16.06-16.86	3.2	1.7	-	-	-
9	139.53-140.63	1.8	-	4.3	5.6	*PTGS2, PLA2G4A*
10	10.58-11.45	1.2	1.5	1.0	6.9	-
12	6.23-7.03	3.4	-	-	-	*DNAI2*

12	40.76-41.56	1.2	1.7	-	-	
13	143.61-144.69	1.2	1.0	-	-	*IQCG*
14	4.13-5.22	1.5	8.7	-	3.7	LOC102167830
14	72.83-73.63	-	-	6.8	-	-
14	99.70-100.51	1.7	1.2	-	-	-
15	61.93-62.73	3.5	-	-	-	-

[a]Chromosome

[b]Position of QTL region

[c]Percentage of genetic variance explained by the QTL region

[d]Semen traits: MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities

[e]The percentage of genetic variance explained by the QTL region is <1%. When the variance is reported for more than one trait, the QTL region is overlapping across traits

[f]Best candidate gene(s) in the region

[g]No candidate genes associated with the trait

Table S2: SNP windows that explained the highest proportion of genetic variance of each semen trait in L1

Trait[a]	Chromosome	Average window position (bp)	Var (%)[b]
MOT	5	18,007,232	7.5
	15	62,333,883	3.5
	12	6,630,173	3.4
PROMOT	14	4,822,304	8.7
	4	28,650,877	6.4
	5	18,074,223	3.0
lnNcells	14	73,227,414	6.8
	1	305,578,277	4.6
	9	139,932,157	4.3
ABN	10	11,041,515	6.9
	1	135,905,888	6.0
	4	28,650,877	5.9

[a]Semen traits: MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities.

[b]Percentage of genetic variance explained by window

Table 4 Individual and overlapping QTL regions for semen traits in L2

Chr[a]	QTL region (Mb)[b]	Var (%)[c] MOT[d]	Var (%) PROMOT	Var (%) lnNcells	Var (%) ABN	Candidate gene[f]
1	270.94-271.74	-[e]	-	9.1	-	-[g]
1	55.61-56.47	2.4	1.4	-	-	-
2	145.69-146.49	-	-	5.2	-	*NME5*
2	154.03-154.83	-	4.3	-	-	-
3	110.29-111.09	-	3.6	-	-	-
6	83.32-84.12	-	-	-	4.7	*AZIN2*
7	116.37-117.28	1.4	-	-	3.3	*SPATA7*
7	82.56-83.36	1.0	2.4	-	-	*METTL3*
8	133.90-134.94	5.9	10.8	-	-	*HPGDS*
9	36.46-37.26	8.0	-	-	-	-
9	9.32-10.31	1.0	1.7	-	-	-
11	41.05-41.85	3.5	1.6	-	-	-
13	11.35-12.15	-	-	-	4.3	-
13	107.48-108.28	8.0	-	-	-	-
15	37.17-37.97	-	-	-	4.9	-
18	42.80-43.60	-	-	4.0	-	-

[a]Chromosome

[b]Position of QTL region

[c]Percentage of genetic variance explained by the QTL region

[d]Semen traits: MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities

[e]The percentage of genetic variance explained by the QTL region is <1%. When the variance is reported for more than one trait, the QTL region is overlapping across traits

[f]Best candidate gene(s) in the region

[g]No candidate genes associated with the trait

Table S3: SNP windows that explained the highest proportion of genetic variance of each semen trait in L2

Trait[a]	Chromosome	Average window position (bp)	Var (%)[b]
MOT	9	36,855,793	8.0
	13	107,880,076	8.0
	8	134,301,890	5.9
PROMOT	8	134,538,185	10.8
	2	154,430,556	4.3
	3	110,686,295	3.6
lnNcells	1	271,338,213	9.1
	2	146,085,202	5.2
	18	43,203,580	4.0
ABN	15	37,574,999	4.9
	6	83,719,373	4.7
	13	11,751,898	4.3

[a]Semen traits: MOT: sperm motility; PROMOT: sperm progressive motility; lnNcells: number of sperm cells per ejaculate; ABN: total morphological abnormalities.

[b]Percentage of genetic variance explained by window

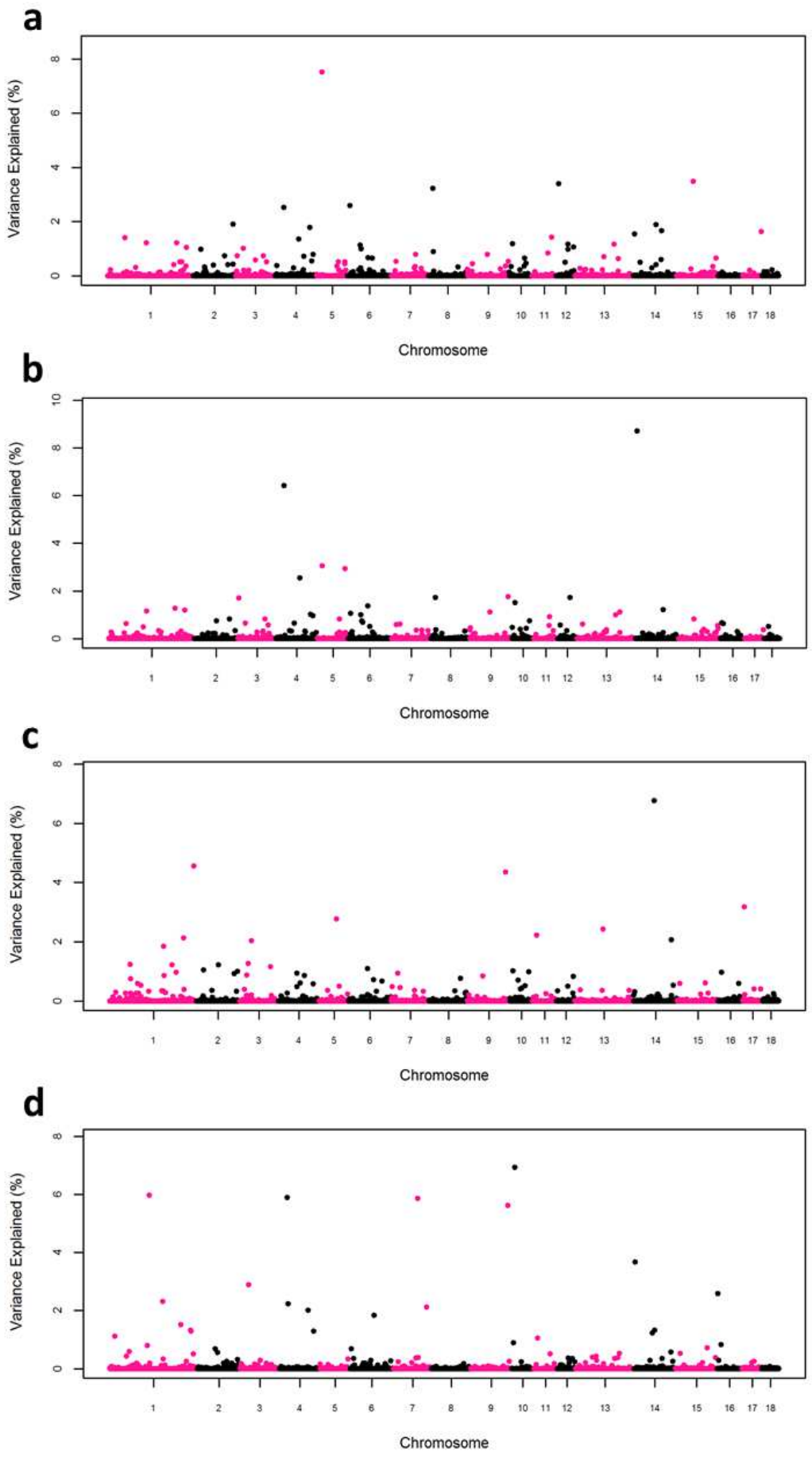
a
Variance Explained (%)
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Chromosome
b
Variance Explained (%)
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Chromosome
c
Variance Explained (%)
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Chromosome
d
Variance Explained (%)
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Chromosome

Fig. 1 GWAS results of semen traits in the Large White type line (L1). a) sperm motility; b) sperm progressive motility; c) number of sperm cells per ejaculate; d) total morphological abnormalities. Each dot represents one SNP window of 0.4 Mb. On y-axis is the percentage of genetic variance explained by windows.

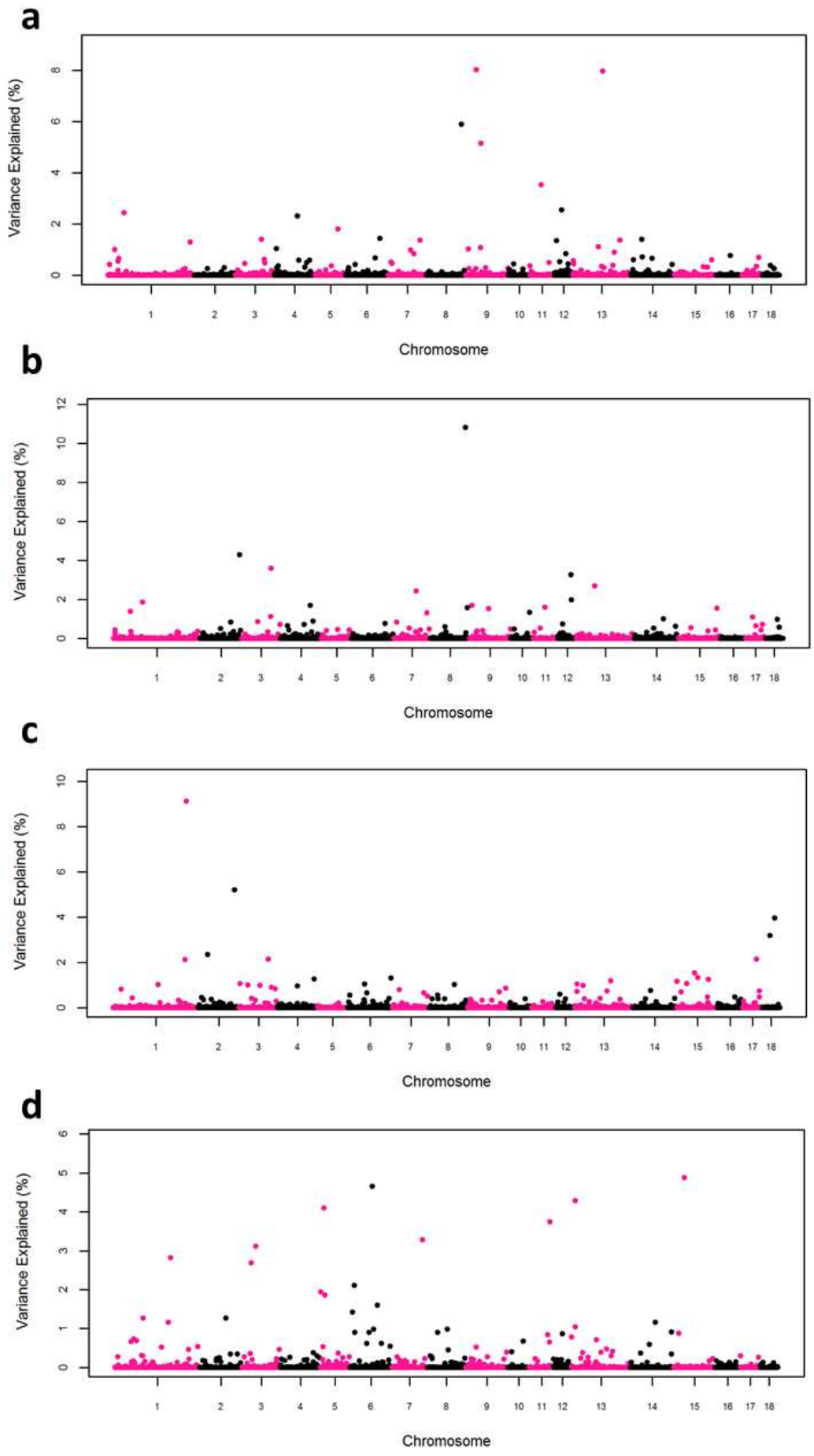
a
Variance Explained (%)
0
2
4
6
8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Chromosome
b
Variance Explained (%)
0
2
4
6
8
10
12
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Chromosome
c
Variance Explained (%)
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Chromosome
d
Variance Explained (%)
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Chromosome

Fig. 2 GWAS results of semen traits in the Landrace type line (L2). a) sperm motility; b) sperm progressive motility; c) number of sperm cells per ejaculate; d) total morphological abnormalities. Each dot represents one SNP window of 0.4 Mb. On y-axis is the percentage of genetic variance explained by windows.

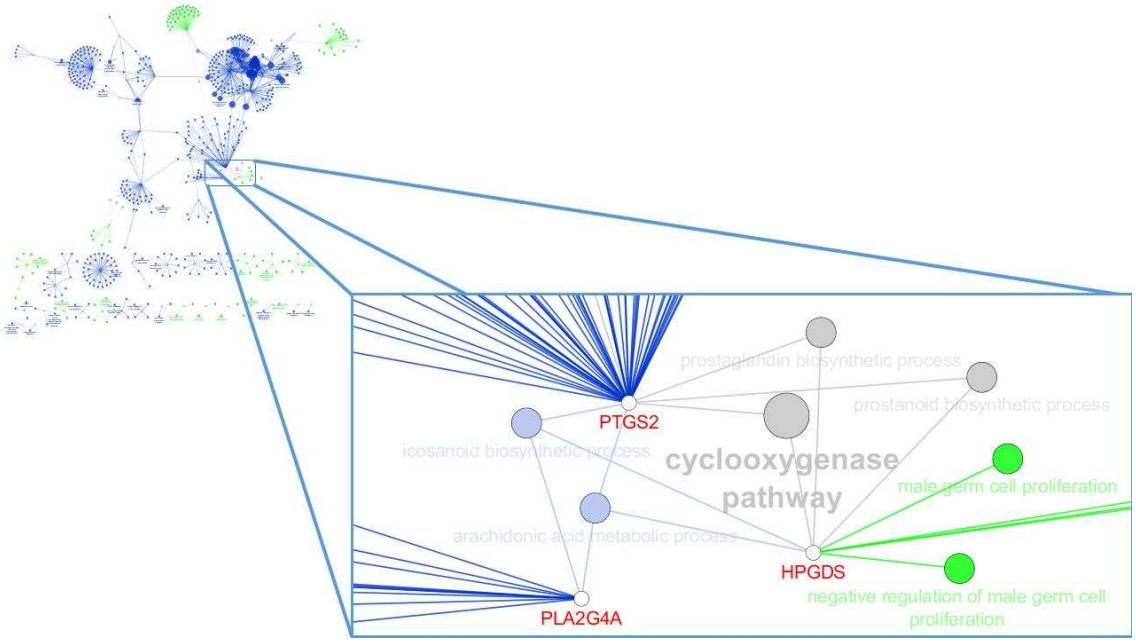

Fig. 3 Gene network of biological processes for progressive motility. Complete network and important shared pathways (with zoom) are shown. Blue color indicates pathways for Large White type line (L1) and green color indicates pathways for Landrace type line (L2). Processes shared by *PTGS2* and *PLA2G4A* genes (L1) and *HPGDS* gene (L2) are connected by blue nodes. Processes shared by *PTGS2* and *HPGDS* are connected by grey nodes. Green dots are biological processes for *HPGDS*.

Discussion

The weighted ssGBLUP approach was chosen to perform the GWAS in this study due to its ability to: 1) integrate all phenotypic, genotypic and pedigree data simultaneously, avoiding the need to calculate pseudo-phenotypes (deregressed EBVs [32]) because of the low number of animals with both genotypes and phenotypes (L1=349, L2=446); 2) allow

different weights for SNPs according to their importance, which deviates from the non-realistic GBLUP statement of infinitesimal model and improves precision of SNP effects estimation; 3) provide possibility to work with SNP windows: considering a window including consecutive SNPs in GWAS could perform better in finding QTL regions compared to individual SNP analysis because of linkage disequilibrium (LD). In general, analyses considering the association of individual SNPs may overestimate the number of detected QTL [33].

In our previous study [7], a single-SNP GWAS was performed for sperm motility in a subset of the L1 population used in the present study. A novel candidate gene (*MTFMT*) associated with MOT on SSC1, position 117.63-117.66 Mb was identified. Regions on SSC1 that overlap between MOT and other semen traits were found in the current study (Table 3) but these regions did not overlap with the *MTFMT* gene described in our previous study [7]. Using the WssGWAS, a higher number of relevant QTL regions were identified for MOT, indicating that this approach was more powerful in detecting QTL. The reasons for not detecting the same QTL region in both studies may be different statistical models used in the studies; the larger number of animals used in the current study and different methodologies (weighted single-step GWAS based on **H** relationship matrix in the current study and single-SNP GWAS using the **A** relationship matrix and deregressed phenotypes in the previous study).

Several relevant QTL regions associated with the traits under study were identified with WssGWAS, confirming the assumption of genetic complexity of these traits. In this study, the criteria used to search for candidate genes considered not only the exact SNP windows, but also their upstream and downstream flanking regions. This is important because high LD may exist between SNPs inside the window and QTL in the surrounding

area. Therefore, a larger genomic region was used for identification of genes. Candidate genes were found for various semen traits in both L1 and L2 pig lines (Table 3 and Table 4). For L1, the dynein axonemal intermediate chain 2 (*DNAI2*) gene, located on SSC12, was considered the best candidate gene for MOT in the QTL region. The *DNAI2* gene (position 6.78-6.80 Mb) encodes axonemal dyneins; the axoneme is a microtubular structure located in the center of all motile cilia and flagella, including sperm flagella [34]. Dyneins are large, multisubunit ATPases that interact with microtubules to generate the driving force for flagellar motility [35]. Mutations in human *DNAI2* were found to be involved in defects in the axoneme [36].

Some overlapping regions were identified between MOT and PROMOT for L1. On SSC5, one candidate gene was identified: the sodium voltage-gated channel alpha subunit 8 (*SCN8A*). Pinto et al. [37] investigated the presence of sodium voltage-gated channels in human spermatozoa and described the expression of *SCN8A* in the sperm flagellum principal piece. They concluded that these sodium channels contribute to the regulation of human sperm motility. Another candidate gene was identified on SSC13: the IQ motif containing G gene (*IQCG*). Li et al. [10] showed that Iqcg knockout mice presented severe malformation and total immobility of their spermatozoa because of disorganized sperm flagellum axoneme. They suggested that Iqcg has a specific role in spermiogenesis. Harris et al. [38] also reported that mice with mutations in Iqcg presented spermiogenesis defects, with incomplete sperm tail formation.

One overlapping region was found for MOT, PROMOT and ABN on SSC14. In this QTL region, the gene LOC102167830, described as spermatogenesis associated protein 31E1-like, is located (4.147-4.154 Mb). In humans, the *SPATA31* is a large gene family with different subfamilies and members, including the *SPATA31E1*. In *mus musculus*, only the

Spata31 gene was described. Wu et al. [39] demonstrated that mouse *Spata31* is localized in the acrosome of round and elongated spermatids. *Spata31* knockout mice showed disorganized testis morphology and aberrant spermatogenic cells in seminiferous tubules. Therefore, this gene was considered as the best candidate gene in this QTL region on SSC14.

The gene network analysis was very useful in investigating shared biological processes among candidate genes identified for the same traits across lines. For PROMOT, we found two genes in L1 (*PLA2G4A* and *PTGS2*) and one gene in L2 (*HPGDS*) that share the biological processes of eicosanoid biosynthesis and arachidonic acid metabolism. Two of the genes, *PTGS2* and *HPGDS*, also share the biological process of cyclooxygenase pathway (Fig. 3 and Fig. 4). Kaewmala et al. [40] investigated the association and expression of *PTGS2*, also known as *COX-2*, as a candidate gene for boar sperm quality and fertility. The presence of *COX-2* protein was described [40] in Leydig cells, spermatogonium and spermatids, suggesting that it may have a role in the spermatogenesis process of boars. They also showed that *COX-2* mRNA and protein tended to be higher in animals with low sperm motility, indicating a negative effect on boar sperm quality. Frungieri et al. [41] showed that *COX-2* enzyme is present in abundance in interstitial cells of men seminiferous tubules with impaired spermatogenesis. The *COX-2* enzymes provide precursor for *HPGDS* action in testes interstitial mast cells [42] (Fig. 4), producing prostaglandin D_2 (PGD_2), which regulate Leydig cell function [42]. Yamamoto et al. [43] and Saharkhiz et al. [44] showed increased sperm motility in men after treatment with mast cells blockers. *HPGDS* is also involved in the negative regulation of male germ cell proliferation (Fig 3), which is linked to PGD_2 production. Moniot et al. [45] identified the PGD_2 pathway as one of the earliest signaling pathway involved in the male germ cell differentiation in the fetal testis.

In the cyclooxygenase pathway, PGH_2 can alternatively be converted in prostaglandin E_2 (PGE_2) and prostaglandin $F_{2\alpha}$ ($PGF_{2\alpha}$) (Fig. 4, IV and V). Schlegel et al. [46] stated that seminal plasma is the richest natural source of prostaglandins, which are synthesized in the seminal vesicles. The authors showed that high concentrations of prostaglandins (especially $PGF_{2\alpha}$) in human seminal fluid was associated with poor sperm motility. Rios et al. [47] demonstrated that low physiological levels of PGE_2 and $PGF_{2\alpha}$ were able to increase and prolong human progressive sperm motility.

Regarding the candidate genes for L2, the NME/NM23 family member 5 gene (*NME5*) was considered as the best candidate gene for lnNcells (Table 4). The *NME5* gene is located on SSC2, position 145.83-145.87 Mb. According to Munier et al. [48], this gene is highly and specifically expressed in testis and the protein encoded by this gene is important for the initial stages of spermatogenesis (before meiosis I). Choi et al. [49] investigated the mechanism by which murine Nm23-M5 (which shares 86% identity with its human homolog *NME5*) protects male haploid germ cells in vivo against oxidative stress-induced apoptosis. They reported that when expression of Nm23-M5 gene is reduced, round and elongated spermatids in testes became more sensitive to oxidative stress, leading to DNA damage and reduced cell numbers, showing that this gene is a critical factor for spermiogenesis. Based on these studies, the NME5 was considered the best candidate gene for lnNcells in this QTL region on SSC2.

The antizyme inhibitor 2 gene (*AZIN2*) is located on SSC6, position 83.49-83.53 Mb, in a region that explained 4.7 % of genetic variance for ABN in L2 (Table 4). Lopez-Contreras [50] showed that mouse *AZIN2* mRNA was exclusively expressed in spermatids. Its expression started in the testis with the beginning of spermiogenesis and the levels of its mRNA increased in more differentiating spermatids, playing a role in redistribution of

polyamines in these cells. The regulation of polyamine levels in haploid germ cells is critical for spermiogenesis [50]. During spermiogenesis, round spermatids elongate, develop an acrosome in sperm head, form a flagellum and dispose excessive cytoplasm [10]. Therefore, if a mutation in *AZIN2* causes impaired spermiogenesis, this may lead to morphological defects in spermatozoa.

The methyltransferase like 3 gene (*METTL3*) is located on SSC7, position 83.00-83.02 Mb, an overlapping region associated with both MOT and PROMOT in L2. According to Liu et al. [51], the *METTL3* and other genes of the methyltransferase complex catalyze the methylation and formation of N^6-methyladenosine (m^6A), which is the most prevalent and reversible RNA epigenetic modification in mammalian mRNA. Yang et al. [52] detected increased m^6A contents in sperm RNA from patients with reduced sperm progressive motility, related to higher expression of *METTL3*. According to these authors, m^6A modification may affect the mRNA expression of genes related to sperm progressive motility. Hence, *METTL3* was considered as the best candidate gene in this SSC7 QTL region.

An overlapping region between MOT and ABN was found in L2 on SSC7 (Table 4). The spermatogenesis associated 7 gene (*SPATA7*) is located in this region (position 116.81-116.86 Mb). SPATA genes are a large gene family that plays a very important role in testis development and spermatogenesis [53]. *SPATA7* was first identified in rat and human spermatocytes and may be involved in preparing chromatin for the initiation of meiotic recombination [54]. According to Ferguson et al. [55], meiotic recombination ties homologous chromosomes, facilitating the proper segregation of chromosomes during meiosis. Problems in crossovers can be responsible for production of aneuploid gametes. Sun et al. [56] prepared a review analyzing the relationship between sperm chromosomal architecture and morphology defects. They concluded that studies analyzing men with sperm

morphological defects and normal sperm concentration show increased frequency of sperm aneuploidies for some chromosomes. Men with a specific type of sperm morphological defect (macrocephalic multiflagellate sperm) showed a high frequency of aneuploid and polyploid sperm. Therefore, the *SPATA7* gene in this QTL region on SSC7 should be considered the best candidate gene for MOT and ABN.

Subsequently, a wider search for candidate genes in QTL regions that explained high proportion of genetic variance (≥5%) was performed for each of the two lines. For that, we extended the search for candidate genes to a larger genomic region (0.8 Mb back and forward from the previously described QTL region in Tables 3 and 4). The ubiquitin specific peptidase 8 (*USP8*) gene is located 0.69 Mb apart from the QTL region for MOT, PROMOT and ABN on SSC1 found in L1 (Table 3). This gene is highly expressed in spermatogenic cells and is involved in the formation and shaping of sperm acrosome and head [57]. Mutations in this gene can lead to acrosome and sperm head defects. Therefore, this gene can be considered a good candidate especially for ABN in this QTL region on SSC1. The spermatogenesis associated 17 (*SPATA17*) gene is located 0.56 Mb apart from the QTL region identified on SSC10 in L1. Nie et al. [58] showed that overexpression of *SPATA17* gene in transgenic mice affected the normal development of male germ cells, with increase in defective spermatogenic cells (apoptotic cells) and giant degenerating cells. Therefore, this gene may be considered a candidate in the region on SSC10. In case of L2, relevant candidate genes in the extended QTL regions were not found.

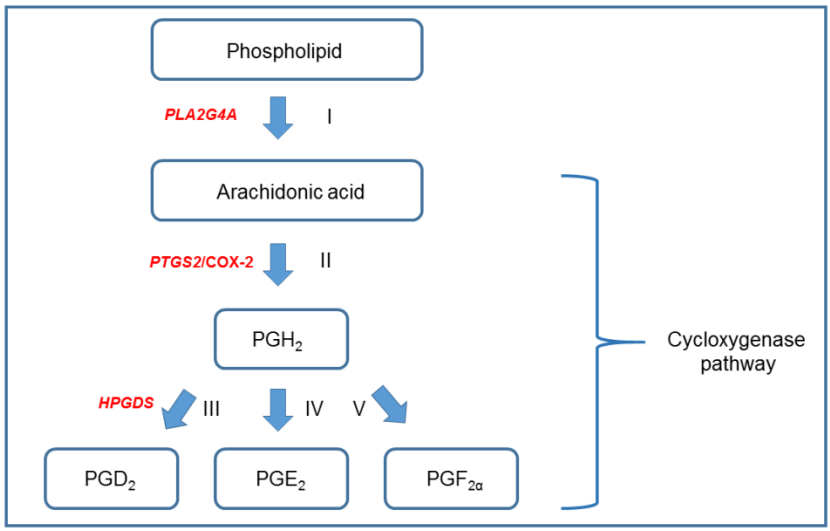

Fig 4. Graphic scheme of pathways shared by genes found in network analysis for progressive motility. Only part of cyclooxygenase pathway is presented. Cytosolic phospholipase A2 group IVA (*PLA2G4A*) is involved in cleaving arachidonic acid from phospholipids, preferentially (I). The free arachidonic acid is then metabolized to produce eicosanoids (including prostaglandins) in the process known as cyclooxygenase pathway (II to V). The genes prostaglandin-endoperoxide synthase 2 (*PTGS2*/COX-2, number II) and hematopoietic prostaglandin D synthase (*HPGDS*, number *III*) are involved in this pathway. The *COX-2* enzymes catalyze prostaglandin H_2 (PGH_2) synthesis from arachidonic acid (II), providing PGH_2 for *HPGDS* action (III) and production of prostaglandin D_2 (PGD_2) in testes interstitial mast cells. PGH_2 can also be converted in prostaglandin E_2 (PGE_2, number IV) and prostaglandin $F_{2\alpha}$ ($PGF_{2\alpha}$, number V).

Conclusions

We identified several QTL regions associated with semen traits in two pig lines using the weighted single-step GWAS. This was possible with low number of animals having both phenotypes and genotypes, due to the appropriate choice of this method. A large part of the genetic variance of the semen traits is explained by different genes in the two lines. Nevertheless, the gene network analysis showed candidate genes sharing biological pathways that occur in mammalian testes, in both the lines as well. These results can be used to search

for causative mutations and for marker-assisted selection to enhance the production and quality of the semen for more effective use of AI in pig breeding and production.

Ethics statement

The data used for this study were obtained as part of routine data recording in a commercial breeding program. Data recording and sample collection were conducted strictly in accordance with Dutch law on the protection of animals (Gezondheids- en welzijnswet voor dieren).

Consent for publication

Not applicable

Data availability

The datasets analysed during the current study are available from Egbert F. Knol (egbert.knol@topigsnorsvin.com) on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

DBDM, MSL, EFK and SEFG conceived the experiment. All authors helped to design the experiment. EFK, MLWJB and MSL provided data for analysis. DBDM performed the analyses. DBDM, JWMB and BH interpreted the results. DBDM drafted the manuscript. All authors improved the writing, read and approved the final manuscript.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Topigs Norsvin.

References

1. Schulze M, Buder S, Rudiger K, Beyerbach M., Waberski D. Influences on semen traits used for selection of young AI boars. Anim Reprod Sci. 2014;148:164-70.

2. Xing Y, Ren J, Ren D, Guo Y, Wu Y, Yang G, et al. A whole genome scanning for quantitative trait loci on traits related to sperm quality and ejaculation in pigs. Anim Reprod Sci. 2009;114:210-8.

3. Sironen A, Uimari P, Nagy S, Paku S, Andersson M, Vilkki J. Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15. BMC Genomics. 2010;11:699.

4. Kaewmala K, Uddin MJ, Cinar MU, Grosse-Brinkhaus C, Jonas E, Tesfaye D, et al. Association study and expression analysis of CD9 as candidate gene for boar sperm quality and fertility traits. Anim Reprod Sci. 2011;125:170-9.

5. Gunawan A, Kaewmala K, Uddin MJ, Cinar MU, Tesfaye D, Phatsara C, et al. Association study and expression analysis of porcine ESR1 as a candidate gene for boar fertility and sperm quality. Anim Reprod Sci. 2011;128:11-21.

6. Gunawan A, Cinar MU, Uddin MJ, Kaewmala K, Tesfaye D, Phatsara C, et al. Investigation on association and expression of ESR2 as a candidate gene for boar sperm quality and fertility. Reprod Domest Anim. 2012;47:782-90.

7. Diniz DB, Lopes MS, Broekhuijse ML, Lopes PS, Harlizius B, Guimaraes SE, et al. A genome-wide association study reveals a novel candidate gene for sperm motility in pigs. Anim Reprod Sci. 2014;151:201-7.

8. Zhao X, Zhao K, Ren J, Zhang F, Jiang C, Hong Y, et al. An imputation-based genome-wide association study on traits related to male reproduction in a White Duroc x Erhualian F2 population. Anim Sci J. 2016;87:646-54.

9. Sarkar H, Arya S, Rai U, Majumdar SS. A study of differential expression of testicular genes in various reproductive phases of *Hemidactylus flaviviridis* (Wall Lizard) to derive their association with onset of spermatogenesis and its relevance to mammals. PLoS One. 2016;11:e0151150.

10. Li RK, Tan JL, Chen LT, Feng JS, Liang WX, Guo XJ, et al. Iqcg is essential for sperm flagellum formation in mice. PLoS One. 2014;9:e98053.

11. Zhang X, Lourenco D, Aguilar I, Legarra A, Misztal I. Weighting strategies for single-Step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. Front Genet. 2016;7:151.

12. Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet Res (Camb). 2012;94:73-83.

13. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci. 2010;93:743-52.

14. Wang H, Misztal I, Aguilar I, Legarra A, Fernando RL, Vitezica Z, et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front Genet. 2014;5:134.

15. Fragomeni Bde O, Misztal I, Lourenco DL, Aguilar I, Okimoto R, Muir WM. Changes in variance explained by top SNP windows over generations for three traits in broiler chicken. Front Genet. 2014;5:332.

16. Lemos MV, Chiaia HL, Berton MP, Feitosa FL, Aboujaoud C, Camargo GM, et al. Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nellore cattle using the single step procedure. BMC Genomics. 2016;17:213.

17. Valente TS, Baldi F, Sant'Anna AC, Albuquerque LG, Paranhos da Costa MJ. Genome-wide association study between single nucleotide polymorphisms and flight speed in Nellore cattle. PLoS One. 2016;11:e0156956.

18. Irano N, de Camargo GM, Costa RB, Terakado AP, Magalhaes AF, Silva RM, et al. Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. PLoS One. 2016;11:e0159502.

19. Tiezzi F, Parker-Gaddis KL, Cole JB, Clay JS, Maltecca C. A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and genomic matrix re-weighting procedure. PLoS One. 2015;10:e0114919.

20. de Oliveira Silva RM, Stafuzza NB, de Oliveira Fragomeni B, de Camargo GMF, Ceacero TM, Cyrillo JNDSG, et al. Genome-wide association study for carcass traits in an experimental Nelore cattle population. PLoS One. 2017;12:e0169860.

21. Verardo LL, Silva FF, Varona L, Resende MD, Bastiaansen JWM, Lopes PS, et al. Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. J Appl Genet. 2015;56:123-32.

22. Verardo LL, Lopes MS, Wijga S, Madsen O, Silva FF, Groenen MA, et al. After genome-wide association studies: gene networks elucidating candidate genes divergences for number of teats across two pig populations. J Anim Sci. 2016;94:1446-58.

23. Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, et al. Analyses of pig genomes provide insight into porcine demography and evolution. Nature. 2012;491:393-98.

24. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81:1084-97.

25. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH: BLUPF90 and related programs (BGF90). In Proceedings of the 7th world congress on genetics applied to livestock production, 19-23 Aug 2002. Montpellier; 2002:28–27.

26. Aguilar I, Misztal I, Legarra A, Tsuruta S. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J Anim Breed Genet. 2011;128:422-8.

27. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91: 4414-23.

28. Legarra A, Robert-Granié C, Manfredi E, Elsen JM. Performance of genomic selection in mice. Genetics. 2008;180:611-8.

29. Veroneze R, Lopes PS, Guimarães SEF, Silva FF, Lopes MS, Harlizius B, et al. Linkage disequilibrium and haplotype block structure in six commercial pig lines. J Anim Sci. 2013;91:3493-501.

30. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25:1091-3.

31. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics. 2013;29:661-3.

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

33. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics. 2011;12:186.

34. McLachlan RI, Ishikawa T, Osianlis T, Robinson P, Merriner DJ, Healy D, et al. Normal live birth after testicular sperm extraction and intracytoplasmic sperm injection in variant primary ciliary dyskinesia with completely immotile sperm and structurally abnormal sperm tails. Fertil Steril. 2012;97:313-8.

35. Witman GB. Axonemal dyneins. Curr Opn Cell Biol. 1992;4:74-9.

36. Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, et al. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. Am J Hum Genet. 2008;83:547-58.

37. Pinto FM, Ravina CG, Fernandez-Sanchez M, Gallardo-Castro M, Cejudo-Román A, Candenas L. Molecular and functional characterization of voltage-gated sodium channels in human sperm. Reprod Biol Endocrinol. 2009;7:71.

38. Harris TP, Schimenti KJ, Munroe RJ, Schimenti JC. IQ motif-containing G (Iqcg) is required for mouse spermiogenesis. G3 (Bethesda). 2014;4:367-72.

39. Wu YY, Yang Y, Xu YD, Yu HL. Targeted disruption of the spermatid-specific gene Spata31 causes male infertility. Mol Reprod Dev. 2015;82:432-40.

40. Kaewmala K, Uddin MJ, Cinar MU, Große -Brinkhaus C, Jonas E, Tesfaye D, et al. Investigation into association and expression of PLCz and COX-2 as candidate genes for boar sperm quality and fertility. Reprod Domest Anim. 2012; 47:213-23.

41. Frungieri MB, Weidinger S, Meineke V, Köhn FM, Mayerhofer A. Proliferative action of mast-cell tryptase is mediated by PAR2, COX2, prostaglandins, and PPARgamma: Possible relevance to human fibrotic disorders. Proc Natl Acad Sci U S A. 2002;99:15072-7

42. Schell C, Frungieri MB, Albrecht M, Gonzalez-Calvar SI, Köhn FM, Calandra RS, et al. A prostaglandin D2 system in the human testis. Fertil Steril. 2007; 88:233-6.

43. Yamamoto M, Hibi H, Miyake K. New treatment of idiopathic severe oligozoospermia with mast cell blocker: results of a single-blind study. Fertil Steril. 1995;64:1221-3.

44. Saharkhiz N, Nikbakht R, Hemadi M. Ketotifen, a mast cell blocker improves sperm motility in asthenospermic infertile men. J Hum Reprod Sci. 2013;6:19-22.

45. Moniot B, Ujjan S, Champagne J, Hirai H, Aritake K, Nagata K, et al. Prostaglandin D2 acts through the Dp2 receptor to influence male germ cell differentiation in the foetal mouse testis. Development. 2014;141:3561-71.

46. Schlegel W, Rotermund S, Färber G, Nieschlag E. The influence of prostaglandins on sperm motility. Prostaglandins. 1981;21:87-99.

47. Rios M, Carreño DV, Oses C, Barrera N, Kerr B, Villalón M. Low physiological levels of prostaglandins E2 and F2alpha improve human sperm functions. Reprod Fertil Dev. 2016;28:434-9.

48. Munier A, Feral C, Milon L, Pinon VP, Gyapay G, Capeau J, et al. A new human nm23 homologue (nm23-H5) specifically expressed in testis germinal cells. FEBS Lett. 1998; 434:289-94.

49. Choi YJ, Cho SK, Hwang KC, Park C, Kim JH, Park SB, et al. Nm23-M5 mediates round and elongated spermatid survival by regulating GPX-5 levels. FEBS Lett. 2009;583:1292-8.

50. López-Contreras AJ, Ramos-Molina B, Martínez-de-la-Torre M, Peñafiel-Verdú C, Puelles L, Cremades A, et al. Expression of antizyme inhibitor 2 in male haploid germinal cells suggests a role in spermiogenesis. Int J Biochem Cell Biol. 2009;41:1070-8.

51. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 2014;10:93-

52. Yang Y, Huang W, Huang JT, Shen F, Xiong J, Yuan EF, et al. Increased N6-methyladenosine in Human Sperm RNA as a Risk Factor for Asthenozoospermia. Sci Rep. 2016;6:24345.

53. Song H, Zhu L, Li Y, Ma C, Guan K, Xia X, et al. Exploiting RNA-sequencing data from the porcine testes to identify the key genes involved in spermatogenesis in Large White pigs. Gene. 2015;573:303-9.

54. Zhang X, Liu H, Zhang Y, Qiao Y, Miao S, Wang L, et al. A novel gene, RSD-3/HSD-3.1, encodes a meiotic-related protein expressed in rat and human testis. J Mol Med (Berl). 2003;81 380-7.

55. Ferguson KA, Wong EC, Chow V, Nigro M, Ma S. Abnormal meiotic recombination in infertile men and its association with sperm aneuploidy. Hum Mol Genet. 2007;16:2870-

56. Sun F, Ko E, Martin RH. Is there a relationship between sperm chromosome abnormalities and sperm morphology?. Reprod Biol Endocrinol. 2006;4:1.

57. Berruti G, Ripolone M, Ceriani M. USP8, a regulator of endosomal sorting, is involved in mouse acrosome biogenesis through interaction with the spermatid ESCRT-0 complex and microtubules. Biol Reprod. 2010;82:930-9.

58. Nie DS, Liu Y, Juan H, Yang X. Overexpression of human SPATA17 protein induces germ cell apoptosis in transgenic male mice. Mol Biol Rep. 2013;40:1905-10.

CHAPTER 4

Effect of service sire and ejaculate on gestation length, total number born and stillborn in pigs[2]

D.B. Diniz,* M.S. Lopes,†‡ M.L.W.J. Broekhuijse,† S.E.F.Guimarães,* E.F. Knol,† R. Veroneze,* J.W.M. Bastiaansen,§ F.F. Silva,* P.S. Lopes*

*Universidade Federal de Viçosa, Animal Science Department, 36570-000, Viçosa-MG, Brazil; †Topigs Norsvin Research Center B.V., P.O. Box 43, 6640 AA Beuningen, the Netherlands; ‡Topigs Norsvin, 80420-210, Curitiba-PR, Brazil; and §Wageningen University, Animal Breeding and Genomics Centre, P.O. Box 338 6700 AH, Wageningen, the Netherlands

[1]The authors thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the financial support. We are also grateful to Topigs Norsvin for providing the data.

Article to be submitted to Journal of Animal Science

ABSTRACT: The efficiency of pig production is highly determined by reproductive traits, such as gestation length (GL), total number of piglets born (TNB) and number of stillborn piglets (SB), which are mainly modeled as traits of the sow, neglecting the paternal effect on these traits. The service sire (boar which ejaculate dose was used to inseminate the sow) and ejaculate effects are two important factors that have the potential to improve the traditional models used in the genetic evaluations of reproductive traits. In this context, we aimed 1) to evaluate the inclusion of genetic service sire and ejaculate effects in the genetic evaluation models of the reproductive traits GL, TNB and log-transformed number of stillborn piglets (lnSB) in a Large White-based commercial pig line and 2) to estimate genetic parameters for service sire on GL, TNB and lnSB. Three single-trait animal models (M1, with service sire included as a permanent environmental effect; M2 with service sire included as both genetic and permanent environmental effects and M3, same as M2, but with inclusion of ejaculate effect) were evaluated. The Akaike Information Criterion (AIC) was applied to select the best model. The M3 showed the smallest AIC values for all traits, indicating that this model fitted better to the data. Applying M3, the heritability estimates for service sire on reproductive traits were low (0.08 ± 0.01 for GL, 0.03 ± 0.01 for TNB and 0.01 ± 0.003 for lnSB), but represented 24.24%, 30% and 11.11%, respectively, of GL, TNB and lnSB heritabilities due to direct (sow) effect. Repeatability estimates for service sire on GL, TNB and lnSB were lower than repeatabilities due to direct effects, showing that permanent environmental effect of sows has a higher influence on the evaluated traits than the permanent environmental effects of service sires. The genetic correlations between direct and service sire effects were positive for all traits (0.75 ± 0.05, 0.18 ± 0.11 and 0.16 ± 0.15 for GL, TNB and lnSB, respectively), which is important because the selection for direct genetic effects

for reproductive traits in sows will affect the paternal component in a positive direction. The results reported in this study indicate that the genetic and permanent environmental effects of service sire on reproductive traits, in addition to the ejaculate effect, should be included in the genetic evaluation models for GL, TNB and lnSB in the considered commercial pig line.

Key words: ejaculate, paternal effect, pig, reproductive traits

INTRODUCTION

The efficiency of pig production is highly determined by reproductive traits, such as gestation length (GL), total number of piglets born (TNB) and stillborn (SB). These traits are commonly included in the selection indexes of pig breeding programs and present low to moderate heritabilities. Therefore, it is important to identify all factors influencing these traits and to include these factors in the genetic evaluation models (Wolf and Wolfová, 2012b). Two of these factors are the paternal effects of service sire (boar which ejaculate dose was used to inseminate the sow) and ejaculate identification. According to Van der Lende et al. (1999), among the factors controlling the litter size, the fertilization rate and prenatal survival rate might be influenced by the service sire due to genetic differences in the capacity of fertilization, which is related to sperm quality (Broekhuijse et al., 2012) and/or the boar genetic contribution to viability of the embryo. Strathe et al. (2013) stated that the genetic variance due to service sire on litter size traits is because of boar fertility and its role in fertilization and embryo survival and eventual piglet survival through the alleles transmitted to the offspring.

Pig breeding programs usually assume only the maternal component of reproductive traits in the genetic evaluation models (Hamann et al., 2004). Other option is the inclusion of additional service sire effect as a non-genetic random effect, which estimates are used for boar culling purposes. Some authors have evaluated the potential for genetic selection of service sire effect on reproductive traits in pigs through variance component and genetic parameter estimations (Van der Lende et al., 1999; Serenius et al., 2003; Hamann et al., 2004; Su et al., 2007; Wolf and Wolfová et al., 2012a). The authors have reported low heritabilities and low genetic correlations between direct and service sire genetic effects. However, no studies have described the effect of the ejaculate used to inseminate the sows on GL, TNB

and SB in pigs. Therefore, our aims were 1) to evaluate the inclusion of genetic service sire and ejaculate effects in the genetic evaluation models for the reproductive traits GL, TNB and log-transformed number of stillborn piglets (lnSB) in a Large White-based commercial pig line and 2) to estimate genetic parameters for service sire on GL, TNB and lnSB.

MATERIALS AND METHODS

Ethics Statement

The data used for this study were obtained as part of routine data recording in a commercial breeding program. Data recording and sample collection were carried out strictly in accordance with the Dutch law on the protection of animals (Gezondheids- en welzijnswet voor dieren).

Animals and reproductive traits

Phenotypic data of GL, TNB and SB were evaluated. The analyzed datasets comprise records from sows inseminated between January 2007 and February 2017. Sows and service sires were from a Large White-based commercial dam line. SB was not normally distributed and was therefore log-transformed (lnSB). The descriptive statistics of data are summarized in Table 1. The dataset for lnSB had the largest number of litters, sows and service sires (47,605, 29,643 and 807, respectively), whereas the dataset for GL had the smallest numbers (33,446, 20,190 and 769, respectively). The pedigree data used in these analysis had a total of 46,483 animals.

Table 1. Descriptive statistics for reproductive traits in a Large White-based pig line

Trait[1]	Litters	Sows	Service sires	Ejaculates	Litters per sow[2]	Litters per service sire[3]	Litters per ejaculate[4]	Mean[5]	Min[6]	Max[7]
GL	33,446	20,190	769	11,575	1.66 (1.06)	43.49 (54.87)	2.89 (2.22)	114.75 (1.48)	105	125
TNB	47,078	29,365	806	14,902	1.60 (1.00)	58.41 (69.72)	3.16 (2.51)	15.45 (3.69)	1	30
lnSB	47,605	29,643	807	15,013	1.61 (1.00)	58.99 (70.52)	3.17 (2.53)	0.63 (0.63)	0	3.14

[1]GL (gestation length), TNB (total number of piglets born), lnSB (log-transformed number of stillborn piglets).

[2]Average number of litters per sow and standard deviation (in parenthesis).

[3]Average number of litters per service sire and standard deviation (in parenthesis).

[4]Average number of litters per ejaculate and standard deviation (in parenthesis).

[5]Mean values for reproductive traits and standard deviation (in parenthesis).

[6]Minimum value for the reproductive trait.

[7]Maximum value for the reproductive trait.

Statistical analysis

Genetic parameters (narrow-sense heritabilities and repeatabilities for direct and service sire effects on GL, TNB and lnSB and genetic correlations between direct and service sire genetic effects) were estimated using three single-trait repeatability animal models implemented in ASReml 3.0 (Gilmour et al., 2009). In model 1 (M1), the service sire effect was included just as a permanent environmental effect. In model 2 (M2), the service sire effect was included as both genetic and permanent environmental effects and the ejaculate effect was not considered. In model 3 (M3) the service sire effect was included as both genetic and permanent environmental effects (as M2), but ejaculate effect was included:

M1: $$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z_1 a} + \boldsymbol{Z_3 p} + \boldsymbol{Z_4 w} + \boldsymbol{\varepsilon}$$

M2: $$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z_1 a} + \boldsymbol{Z_2 s} + \boldsymbol{Z_3 p} + \boldsymbol{Z_4 w} + \boldsymbol{\varepsilon}$$

M3: $$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z_1 a} + \boldsymbol{Z_2 s} + \boldsymbol{Z_3 p} + \boldsymbol{Z_4 w} + \boldsymbol{Z_5 e} + \boldsymbol{\varepsilon}$$

in which $\boldsymbol{y}$ is the vector of phenotypic observations (GL, TNB or lnSB); $\boldsymbol{\beta}$ is the vector of fixed effects (herd-year-season of farrowing; repetition or not of insemination (within the same cycle); the linear covariate of parity number; the class of interval between weaning to pregnancy for TNB and lnSB); $\boldsymbol{a}$ is the vector of direct additive genetic effects; $\boldsymbol{s}$ is the vector of service sire additive genetic effects; $\boldsymbol{p}$ is the vector of permanent environmental effects of sow; $\boldsymbol{w}$ is the vector of permanent environmental effects of service sire; $\boldsymbol{e}$ is the vector of ejaculate effects nested within service sire; $\boldsymbol{\varepsilon}$ is the vector of random residuals; and $\mathbf{X}$, $\boldsymbol{Z_1}$, $\boldsymbol{Z_2}$, $\boldsymbol{Z_3}$, $\boldsymbol{Z_4}$ and $\boldsymbol{Z_5}$ are the incidence matrices of $\boldsymbol{\beta}$, $\boldsymbol{a}$, $\boldsymbol{s}$, $\boldsymbol{p}$, $\boldsymbol{w}$ and $\boldsymbol{e}$, respectively.

For M1, M2 and M3, the random effects were assumed to be normally distributed with mean zero and covariance structure given by, respectively:

$$Var\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{p}\\ \boldsymbol{w}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}\sigma_a^2 & 0 & 0 & 0\\ 0 & \mathbf{I}\sigma_p^2 & 0 & 0\\ 0 & 0 & \mathbf{I}\sigma_w^2 & 0\\ 0 & 0 & 0 & \mathbf{I}\sigma_{\boldsymbol{\varepsilon}}^2\end{bmatrix},$$

$$Var\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{s}\\ \boldsymbol{p}\\ \boldsymbol{w}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}\sigma_a^2 & \boldsymbol{A}\sigma_{as} & 0 & 0 & 0\\ \boldsymbol{A}\sigma_{as} & \mathbf{A}\sigma_s^2 & 0 & 0 & 0\\ 0 & 0 & \mathbf{I}\sigma_p^2 & 0 & 0\\ 0 & 0 & 0 & \mathbf{I}\sigma_w^2 & 0\\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\boldsymbol{\varepsilon}}^2\end{bmatrix} \text{ and}$$

$$Var\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{s}\\ \boldsymbol{p}\\ \boldsymbol{w}\\ \boldsymbol{e}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}\sigma_a^2 & \boldsymbol{A}\sigma_{as} & 0 & 0 & 0 & 0\\ \boldsymbol{A}\sigma_{as} & \mathbf{A}\sigma_s^2 & 0 & 0 & 0 & 0\\ 0 & 0 & \mathbf{I}\sigma_p^2 & 0 & 0 & 0\\ 0 & 0 & 0 & \mathbf{I}\sigma_w^2 & 0 & 0\\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_e^2 & 0\\ 0 & 0 & 0 & 0 & 0 & \mathbf{I}\sigma_{\boldsymbol{\varepsilon}}^2\end{bmatrix},$$

in which σ_a^2, σ_s^2, σ_p^2, σ_w^2, σ_e^2, $\sigma_{\boldsymbol{\varepsilon}}^2$ and σ_{as} represent the variances due to direct additive genetic effect, service sire additive genetic effect, permanent environmental effect of sow, permanent environmental effect of service sire, ejaculate effect, residual variance and covariance between direct and service sire genetic effects, respectively. The pedigree-based relationship matrix is defined by **A**; and **I** is an identity matrix.

Narrow-sense heritabilities for direct additive genetic effect (h_a^2) on GL, TNB and lnSB were estimated for M1, M2 and M3 as:

$h_a^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_p^2 + \sigma_w^2 + \sigma_\varepsilon^2)$,

$h_a^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_{as} + \sigma_\varepsilon^2)$ and

$h_a^2 = \sigma_a^2 / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_e^2 + \sigma_{as} + \sigma_\varepsilon^2)$,

in which σ_a^2, σ_p^2, σ_s^2, σ_w^2, σ_e^2, σ_{as} and σ_ε^2 are as defined before.

Narrow-sense heritabilities for service sire additive genetic effect (h_s^2) on GL, TNB and lnSB were estimated for M2 and M3 as:

$h_s^2 = \sigma_s^2 / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_{as} + \sigma_\varepsilon^2)$ and

$h_s^2 = \sigma_s^2 / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_e^2 + \sigma_{as} + \sigma_\varepsilon^2)$

Repeatabilities for direct effect (t_a) on GL, TNB and lnSB were estimated for M1, M2 and M3 as:

$t_a = (\sigma_a^2 + \sigma_p^2) / (\sigma_a^2 + \sigma_p^2 + \sigma_w^2 + \sigma_\varepsilon^2)$,

$t_a = (\sigma_a^2 + \sigma_p^2) / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_{as} + \sigma_\varepsilon^2)$ and

$t_a = (\sigma_a^2 + \sigma_p^2) / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_e^2 + \sigma_{as} + \sigma_\varepsilon^2)$

Repeatabilities for service sire effect (t_s) on GL, TNB and lnSB were estimated for M2 and M3 as:

$t_s = (\sigma_s^2 + \sigma_w^2) / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_{as} + \sigma_\varepsilon^2)$ and

$$t_s = (\sigma_s^2 + \sigma_w^2 + \sigma_e^2) / (\sigma_a^2 + \sigma_p^2 + \sigma_s^2 + \sigma_w^2 + \sigma_e^2 + \sigma_{as} + \sigma_\varepsilon^2)$$

Genetic correlations between direct and service sire genetic effects were estimated for M2 and M3 as:

$$r_{as} = \sigma_{as} / \sqrt{\sigma_a^2 * \sigma_s^2}$$

The Akaike Information Criterion (AIC) (Akaike, 1973) was used to select the best model. The AIC was calculated as follows (Gilmour et al., 2009):

$$AIC = -2LogL + 2t_i,$$

in which $LogL$ is the maximum REML log-likelihood of the model and t_i is the number of variance parameters in the model. Model with the smallest value of AIC is preferred.

RESULTS

For all traits, the AIC values were smaller for M3, revealing that the model with inclusion of the service sire as genetic and permanent environmental effects, in addition to the ejaculate effect, presented the better fit to the data. Therefore, all the results and implications will be reported and discussed based on M3.

Variance components, genetic parameter estimates and AIC values for the reproductive traits provided by each model used are shown in Table 2. Variance estimates were very similar for all traits, except for permanent environmental variances due to service sire, which were greater in M1 and residual variances, which were lower in M3. In this case, with inclusion of ejaculate effect in M3, the residual variances decreased and the repeatabilities due to service sire effect slightly increased. Repeatabilities due to direct effect did not change among

models, except for GL, which presented greater repeatability in M1. Repeatabilities due to service sire effect were lower than repeatabilities due to direct effect for all traits. GL showed greater repeatabilities due to direct effect (0.40 ± 0.01) compared to TNB and lnSB (0.21 ± 0.01 and 0.17 ± 0.01, respectively).

Heritabilities for direct and service sire effects on TNB and lnSB did not change, whereas on GL, heritability for direct effect was greater in M1. The estimated heritability for direct effect on GL was moderate (0.33 ± 0.01) and on TNB and lnSB they were low (0.10 ± 0.01 and 0.09 ± 0.01, respectively). In general, the estimated heritabilities for service sire effect on GL, TNB and lnSB were low (0.08 ± 0.01, 0.03 ± 0.01 and 0.01 ± 0.003, respectively) and represented 24.24%, 30% and 11.11% of heritabilities for direct effects on GL, TNB and lnSB, respectively.

The genetic correlations between direct and service sire effects were positive for all traits. A greater genetic correlation was found for GL (0.75 ± 0.05) compared to TNB and lnSB (0.18 ± 0.11 and 0.16 ± 0.15, respectively).

Table 2. Variances due to direct additive genetic effect (σ_a^2), service sire additive genetic effect (σ_s^2), permanent environmental effect of sow (σ_p^2), permanent environmental effect of service sire (σ_w^2) and ejaculate effect ($\sigma_{\mathbf{e}}^2$); residual ($\sigma_{\boldsymbol{\varepsilon}}^2$) and phenotypic ($\sigma_{\mathbf{t}}^2$) variances; covariance between direct and service sire genetic effects (σ_{as}); repeatabilities for direct (t_a) and service sire (t_s) effects; heritabilities for direct (h_a^2) and service sire (h_s^2) effects; genetic correlations between direct and service sire genetic effects (r_{as}).

Model[1]	σ_a^2	σ_s^2	σ_p^2	σ_w^2	$\sigma_{\mathbf{e}}^2$	$\sigma_{\boldsymbol{\varepsilon}}^2$	$\sigma_{\mathbf{t}}^2$	σ_{as}	t_a	t_s	h_a^2	h_s^2	r_{as}	AIC[3]
Gestation length														
1	0.75 (0.03)[2]	-	0.17 (0.02)	0.19 (0.01)	-	0.89 (0.01)	2.00 (0.03)	-	0.46 (0.01)	-	0.37 (0.01)	-	-	8009.46
2	0.74 (0.03)	0.17 (0.02)	0.17 (0.02)	0.03 (0.01)	-	0.89 (0.01)	2.28 (0.05)	0.27 (0.02)	0.40 (0.01)	0.09 (0.01)	0.33 (0.01)	0.08 (0.01)	0.74 (0.05)	7911.96

3	0.74 (0.03)	0.17 (0.02)	0.17 (0.02)	0.02 (0.01)	0.06 (0.01)	0.84 (0.01)	2.28 (0.05)	0.27 (0.02)	0.40 (0.01)	0.11 (0.01)	0.33 (0.01)	0.08 (0.01)	0.75 (0.05)	7860.02
Total number of piglets born														
1	1.17 (0.11)	-	1.36 (0.11)	0.49 (0.05)	-	8.88 (0.10)	11.90 (0.10)	-	0.21 (0.01)	-	0.10 (0.01)	-	-	10537.40
2	1.17 (0.11)	0.32 (0.09)	1.36 (0.11)	0.19 (0.07)	-	8.88 (0.10)	12.03 (0.13)	0.10 (0.07)	0.21 (0.01)	0.04 (0.004)	0.10 (0.01)	0.03 (0.01)	0.17 (0.11)	10518.00
3	1.16 (0.11)	0.32 (0.09)	1.37 (0.11)	0.18 (0.07)	0.20 (0.05)	8.72 (0.10)	12.05 (0.13)	0.11 (0.07)	0.21 (0.01)	0.06 (0.01)	0.10 (0.01)	0.03 (0.01)	0.18 (0.11)	10502.70
log-transformed number of stillborn piglets														
1	0.032 (0.003)	-	0.029 (0.003)	0.004 (0.001)	-	0.297 (0.003)	0.362 (0.003)	-	0.17 (0.01)	-	0.09 (0.01)	-	-	8922.16
2	0.032 (0.003)	0.003 (0.001)	0.029 (0.003)	0.002 (0.001)	-	0.297 (0.003)	0.364 (0.003)	0.001 (0.001)	0.17 (0.01)	0.01 (0.002)	0.09 (0.01)	0.01 (0.003)	0.15 (0.15)	8911.46
3	0.032 (0.003)	0.003 (0.001)	0.029 (0.003)	0.001 (0.001)	0.003 (0.001)	0.294 (0.003)	0.364 (0.003)	0.001 (0.001)	0.17 (0.01)	0.02 (0.004)	0.09 (0.01)	0.01 (0.003)	0.16 (0.15)	8908.74

[1]Model 1: service sire as permanent environmental effect; model 2: service sire as genetic and environmental effect, without inclusion of ejaculate effect; model 3: service sire as genetic and environmental effect, with inclusion of ejaculate effect.

[2]Standard errors are given in parenthesis.

[3]Akaike Information Criterion.

DISCUSSION

Reproductive traits, such as GL, TNB and SB, in addition with other important production and reproductive traits, are the bottleneck for efficient pig production. Genetic studies for these traits have been extensively performed and their selection in pig breeding programs is a routine for many years. However, the paternal side of reproductive traits is continuously neglected. This paternal effect is given by both genetic and environmental components. Frequently, when the service sire effect is accounted for in pig breeding selection indexes, it is included only as a permanent environmental effect, whose estimates are used for boar culling purposes. Nevertheless, the decision to include the service sire as genetic or its environmental effect in statistical models of reproductive traits is still not a

consensus in the literature. One reason for that may be the low heritability for the paternal effect.

Van der Lende et al. (1999) evaluated the service sire effect on TNB and number of piglets born alive (NBA) in 36,708 litters from 1,044 service sires. Heritability estimates for service sire effect on TNB and NBA ranged from 0.002 to 0.04 and the authors concluded that genetic selection for service sire effect would not improve litter size in pigs. Serenius et al. (2003) analyzed data containing litter records of 6,514 Finnish Large White and 9,154 Landrace pigs. Heritabilities and repeatabilities estimates for service sire effect on TNB ranged from 0.008 to 0.012 and from 0.013 to 0.030, respectively. The authors concluded that besides the low effect of service sire, it should be considered in the statistical model as a random environmental effect for culling of boars with poor piglet production. Wolf and Wolfová (2012a) evaluated the inclusion of service sire as random or genetic effect in genetic evaluation models for TNB, NBA and number of piglets weaned. The authors estimated proportion of variance for service sire and heritabilities for service sire on TNB, NBA and number of piglets weaned in the range from 2.0 to 3.0% and 0.02 and 0.03, respectively, and concluded that models without service sire effect or models including service sire just as random effect are proposed for genetic evaluation of litter size traits. Wolf and Wolfová (2012b) reported very low proportion of variance due to service sire (between 0.8 and 1.6 %) for SB and piglets mortality between 24 hours after birth and weaning (service sire included in the model only as a random non-genetic effect), and concluded that it is not necessary to include the service sire in statistical models for these traits. Su et al. (2007) reported repeatability of sire for litter size ranging from 0.040 to 0.065 and service sire heritability estimates on TNB ranging between 0.02 and 0.03. They concluded that it is reasonable to include the service sire effect in a model for genetic evaluation of litter size.

In the present study, we found additive genetic variances due to service sire effect of 0.17 ± 0.02, 0.32 ± 0.09 and 0.003 ± 0.001, which represents the proportion of 22.97%, 27.59% and 9.38% of additive genetic variance due to direct effect for GL, TNB and lnSB, respectively. In literature, the proportion between additive genetic variances due to service sire effect and additive genetic variance due to direct effect on TNB ranged from 2.12% to 39.39% (Van der Lende et al., 1999; Kim et al., 2002; Serenius et al., 2003; Su et al., 2007; Wolf and Wolfová et al., 2012a; Strathe et al., 2013). For SB, the service sire was evaluated only as a non-genetic effect (Holm et al., 2004; Wolf and Wolfová et al., 2012b) and there are no reports in literature evaluating the service sire effect on GL. Our results show that, as expected, the additive genetic effect due to sow is the main genetic component influencing the evaluated reproductive traits, but the genetic of service sire should also be considered, especially for GL and TNB, because of the greater proportion between additive genetic variances due to service sire and direct effects for these traits. According to Hamman et al. (2004), the genetic variance for NBA due to service sire effect in their study might be due to genetic differences in quality and quantity of semen or to genetic effects controlling fertilization of oocytes and survival of embryos and foetuses, as the chromosomal translocations. The service sires used in the present study have been evaluated for semen quality and quantity and also checked for chromosomal translocations. Therefore, these would not be the reason for the genetic variation due to service sire effect on GL, TNB and lnSB. Other genetic factors, for example, copy number variations (CNVs), which has recently been studied involved in boar fertility (Revay et al., 2015), could be a possible reason for the reported genetic differences between service sires. The lower genetic variance for lnSB due to service sire may be explained by the minor importance of paternal effects on later stages

of fetal development, when embryos become more dependent on uterine environment and maternal influences are predominant (Hamman et al., 2004).

In our study, a model with inclusion of service sire just as a permanent environmental effect (M1) was compared with models in which the service sire was included as both genetic and permanent environmental effects (M2 and M3). Estimates for permanent environmental variances due to service sire were greater in M1 and residual variances were lower in M3 for all traits, whereas the other variance component estimates were very similar. Greater variances due to permanent environmental effect of service sire were estimated in M1 because the portion of phenotypic variance that would be explained by the genetic effect of service sire was accounted by the permanent effect in M1 (the sum of permanent environmental and genetic variances due to service sire effect in M3 is equal to the estimates of permanent environmental variances due to service sire in M1, approximately, for all traits) (Table 2). The residual variances decreased with inclusion of ejaculate effect in M3 because the ejaculate effect took a portion of unexplained variance, which was in the residual effect in M1 and M2.

Repeatabilities due to service sire effects slightly increased in M3 because of the additional component of ejaculate variance in this model. Greater repeatabilities and heritabilities due to direct effects were found for GL in M1 because in the other models, there is the addition of the genetic covariance between direct and service sire effects on phenotypic variance, which was not present in M1, increasing the repeatabilities and heritabilities. For the other traits, the repeatabilities and heritabilities due to direct effects did not change among models because the proportions between the estimates of covariance between direct and service sire effects and phenotypic variances were lower compared to GL. There are evident permanent environmental effects due to direct and service sire effects influencing all the

reproductive traits, since the repeatability estimates were greater than the heritability estimates. Because repeatabilities due to service sire effect were lower than repeatabilities due to direct effect for all traits, the permanent environmental effect of sows influence more GL, TNB and lnSB than permanent environmental effects due to service sire. The greater repeatability for GL due to direct effect (0.40 ± 0.01) indicates that, in any parity, sows were more likely to have similar GL to a previous GL, which helps swine producers to be ready to assist farrowing, or to guide the decisions, as to induce parturition (Sasaki and Koketsu, 2007).

Low heritability for service sire effect on TNB (0.03 ± 0.01) is in agreement with other reports in literature, which ranged from 0.002 to 0.03 (Van der Lende et al., 1999; Kim et al., 2002; Serenius et al., 2003; Su et al., 2007; Wolf and Wolfová et al., 2012a; Strathe et al., 2013). Although low, the heritability estimates varied, which can be explained by different models applied (with or without inclusion of permanent environmental effects due to service sire; considering or not the covariance between direct and service sire effects), diverse amount of data and different breeds and crosses used.

Moderate to high (0.75 ± 0.05) and low (0.18 ± 0.11 and 0.16 ± 0.15) genetic correlations between direct and service sire effects were estimated for GL, TNB and lnSB, respectively. For TNB, Van der Lende et al. (1999) reported genetic correlations between direct and service sire genetic effects ranging from -0.47 to 0.08 in purebred pig lines. Serenius et al. (2003) estimated genetic correlations of 0.23 ± 0.21 and 0.21 ± 0.13 for the first and other parities, respectively, in a Finish Large White pig breed. In Wolf and Wolfová et al. (2012a), the estimated genetic correlation was –0.24 ± 0.22 for a Czech Large White breed. As indicated by Serenius et al. (2003) and Wolf and Wolfová et al. (2012a), the high standard errors of estimates (also revealed in our study mainly for lnSB) express unstable and no clear

relationship between direct and service sire genetic effects. Nevertheless, all estimates were positive, which is important because the selection for direct genetic effects for reproductive traits in sows will affect the paternal component in a favorable direction.

The inclusion of ejaculate effect in the statistical models for reproductive traits could be important to allow the estimation of the effects of different ejaculates of the same service sire on GL, TNB and lnSB and help to select the boars with optimal estimated effects (intermediate for GL, greater for TNB and lower for lnSB). The AIC values were lower for M3, indicating that the best model was the one which included the ejaculate effect. To our knowledge, this is the first study in literature that evaluated the inclusion of ejaculate effect in genetic models for reproductive traits in pigs.

Although most studies in literature evaluating the service sire effect on reproductive traits in pigs state that it is not necessary to include this effect in the genetic evaluation models for these traits or it is recommended to include the service sire just as a random non-genetic effect (Van der Lende et al., 1999; Serenius et al., 2003; Su et al., 2007; Wolf and Wolfová et al., 2012a,b), in the present study we verified that there is genetic variation due to service sire genetic effect on GL, TNB and lnSB. In addition, this variation accounts for 22.97%, 27.59% and 9.38% of additive genetic variance due to direct effect for GL, TNB and lnSB, respectively. Moreover, the model with inclusion of permanent environmental and genetic effects due to service sire showed the best fit to the data. These results might indicate that the genetic and permanent environmental effects due to service sire, in addition to the ejaculate effect, should be included in the genetic evaluation models for GL, TNB and lnSB in the considered commercial pig line. Service sire estimated breeding values (EBVs) for reproductive traits would be more suitable for breeding purposes and culling of boars than only permanent environmental estimates. In order to achieve more evident and faster

improvement of boar fertility, evaluation of ejaculate effect and estimation of genetic parameters for reproductive traits due to service sire effect should also be performed in pig sire lines, because of the use of boars directly in the three way cross schemes widely used in breeding programs.

LITERATURE CITED

Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory (eds B.N. Petrov & F. Csaki), pp. 267–281. Akademiai Kiado, Budapest.

Broekhuijse, M. L. W. J., E. Šoštarić, H. Feitsma, and B. M. Gadella. 2012. Application of computer-assisted semen analysis to explain variations in pig fertility. J. Anim. Sci. 90: 779-789. doi: 10.2527/jas.2011-4311

Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, and D. Butler. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.

Hamann, H., R. Steinheuer, and O. Distl. 2004. Estimation of genetic parameters for litter size as a sow and boar trait in German herdbook Landrace and Pietrain swine. Livest. Prod. Sci.85:201–207. doi: 10.1016/S0301-6226(03)00135-0.

Holm B., M. Bakken, O.Vangen, and R. Rekaya. 2004. Genetic analysis of litter size, parturition length, and birth assistance requirements in primiparous sows using a joint linear-threshold animal model. J. Anim. Sci. 82: 2528–2533.

Kim B.W., S. D. Kim, I. J. Lee, K. H. Chung, O. S. Kwon, J. K. Ha, and J. G. Lee. 2002. Estimation of direct and service sire genetic parameters for reproductive traits in Yorkshire. Asian-Australas. J. Anim Sci. 15: 1232–1236. doi: 10.5713/ajas.2002.1232.

Revay T., A. T. Quach, L. Maignel, B. Sullivan, and W. A. King. 2015. Copy number variations in high and low fertility breeding boars. BMC Genomics 16: 280. doi: 10.1186/s12864-015-1473-9.

Sasaki Y., and Y. Koketsu. 2007. Variability and repeatability in gestation length related to litter performance in female pigs on commercial farms. Theriogenology. 68:123–127. doi: 10.1016/j.theriogenology.2007.04.021.

Serenius, T., M.-L. Sevón-Aimonen, and E. A. Mantysaari. 2003. Effect of service sire and validity of repeatability model in litter size and farrowing interval of Finnish Landrace and Large White populations. Livest. Prod. Sci. 81:213–222. doi: 10.1016/S0301-6226(02)00300-7.

Strathe, A. B., I. H. Velander, T. Mark, T. Ostersen, C. Hansen, and H. N. Kadarmideen. 2013. Genetic parameters for male fertility and its relationship to skatole and androstenone in Danish Landrace boars. J Anim Sci. 91:4659–4668. doi: 10.2527/jas.2013-6454.

Su G., M. S. Lund, and Sorensen D. 2007. Selection for litter size at day five to improve litter size at weaning and piglet survival rate. J. Anim. Sci., 85: 1385–1392. doi: 10.2527/jas.2006-631.

Van der Lende, T., M. H. A. Willemsen, J. A. M. Van Arendonk, and E. B. P. G. Van Haandel. 1999. Genetic analysis of the service sire effect on litter size in swine. Livest. Prod. Sci. 58:91–94. doi: 10.1016/S0301-6226(98)00182-1.

Wolf J., and M. Wolfová. 2012a. Effect of service sire on litter size traits in Czech Large White and Landrace pigs. Czech J Anim Sci. 57: 220-230.

Wolf J., and M. Wolfová. 2012b. Genetic parameters including the service sire effect for the sow traits stillbirth and piglet losses in Czech Large White and Landrace. Czech J Anim Sci 57: 402-409.

CHAPTER 5

GENERAL DISCUSSION

Introduction

Increasing development in pig breeding industry requires studies and evaluation of market demand in order to maximize the genetic progress. Following this tendency, new important traits, which until the moment are not considered in breeding goals, are being evaluated for their inclusion in selection indexes. Boar semen and fertility traits are examples of these traits, which are extremely important in pig production and dissemination of genetic progress.

In this thesis, I evaluated the genetic parameters for semen quality and quantity traits in five different pig dam and sire lines (Chapter 2). Although dam lines are specialized in reproductive traits such as litter size and gestation length, and sire lines are specialized in production traits, boars from both lines must produce semen in sufficient quantity and quality to surpass the quality threshold to be used for production of insemination doses. The selection of boars with lower variability in semen production and quality during their reproductive life is also important to reduce the discard of ejaculates that do not meet the threshold criteria. In the establishment of a breeding program, there are several important steps to be followed and the estimation of genetic parameters for target traits is crucial in the long-term. After the definition of the production system, choice of target traits, definition of breeding goals and collection of phenotypes, genotypes and family information, the knowledge of genetic parameters becomes important and necessary.

Evolving in the genetic analysis of boar semen traits, in Chapter 3, I studied genomic regions and candidate genes for the same semen traits studied in the previous chapter. Genome-wide association studies (GWAS) were applied, which are useful tools to aid the elucidation of the genetic control of quantitative traits and relevant genetic markers pinpointed in GWAS can be used for marker-assisted selection in the evaluated pig lines. The study of boar reproduction was extended to the fertility field in Chapter 4, in which the effects of service sire and ejaculate were evaluated for reproductive traits that are commonly treated as maternal traits, like gestation length (GL), total number of piglets born (TNB) and stillborn (SB). Genetic parameters and variation for service sire on these traits were also

estimated to support the decision to include the genetic paternal side on the genetic evaluation models for reproductive traits. In each of the chapters, the main results have already been discussed.

In this discussion chapter, I will give a broader explanation of the importance to genetically evaluate the boar semen and fertility traits, starting with a briefly description of the pig breeding structure, proceeding with the importance of the artificial insemination and semen quality in this context and of new genomic technologies that support the increasing knowledge of the genes controlling the semen traits. Finally, I will discuss the main aspects of the boar fertility field.

The pig breeding industry

The pig breeding industry is a well-organized sector that has its traditional pyramidal structure composed by three levels: nucleus, multiplier and commercial. In the top of the pyramid are the nucleus farms, which represent smaller number of animals and have different and specialized pure and synthetic dam and sire lines (Lopes, 2016). A small number of multinational breeding companies controls most of the nucleus farms, which invest in genetic and genomic research and apply the selection based on diverse breeding goals depending on the consumer demand; in the middle of the pyramid are the multipliers, who receive animals from nucleus farms and perform the crosses. In the basis of the pyramid and in higher number are the commercial farms, where the terminal crossbred animals are produced to be sold to consumer market. The majority of studies about breeding and genetics for all the purebred lines and traits are applied first in the nucleus farms, and, consequently, they are passed downwards to inferior levels through improved animals and/or semen. The crosses performed in multipliers and commercial levels are designed to capture the advantages of heterosis and complementarity of breeds (Knox, 2016).

Years ago, when pig genetic improvement started to be applied in practice in an industrial context, the breeding goal was focused essentially on production traits, such as growth and backfat thickness. Then, over the years, traits with lower heritabilities, as reproductive traits (mainly litter size), started to be added in the breeding objectives (Merks, 2000; Lopes, 2016). The application of the Best Linear Unbiased Prediction (BLUP; Henderson, 1963, 1975), the advance in computational tools and statistical methods, and improvements in performance recording enabled the inclusion of other traits in the selection

indexes of breeding companies (Merks, 2000; Merks et al., 2012). In practice, breeding programs combine all economically important traits in a single index, which is the basis for breeding values estimation and selection of animals (Hermesch et al., 2015).

The importance of artificial insemination and evaluation of semen traits for genetic progress

The introduction of artificial insemination (AI) in pig production and breeding industries was a very important development, which enabled improved health control, dissemination of genetic progress, increased selection pressure and a better profitability of each boar ejaculate (Rodriguez et al., 2017). In the pig industry, with the use of AI, a single male has a higher impact on efficiency and productivity than a female (Ruiz-Sánchez et al., 2006).

Boars selected for AI are genetically evaluated for economically important traits, as growth rate and fatness (Robinson and Buhr, 2005). In addition, boars are evaluated for additional factors affecting their suitability for AI, as conformation defects, chromosomal translocations and semen quality (Robinson and Buhr, 2005; Maes et al., 2011). However, according to Robinson and Buhr (2005), breeding values for commercially important traits must not be the exclusive criteria for boar selection, since animals with poor semen production may be selected. The fertility standard reached with AI, considering the variation among countries and farms, can be attributed to the technology for sperm production and quality monitoring (Knox, 2016).

Different semen traits can be used to assess semen quality. Among them, motility and progressive motility, which are the proportion of moving sperm cells in the ejaculate and proportion of sperm cells moving in a straight line, respectively, are the most common traits used to evaluate que semen quality (Gadea, 2005). Such importance may be related to their relationship with fertilizing capacity (Diniz et al., 2014). In addition, the assessment of morphological abnormalities (defects in sperm cells) in spermatozoa is also crucial for good standard of boar fertility (Alm et al., 2006; Bracket, 2006). Sperm with morphological defects, especially in the tail, would have mechanical problems to generate a normal flagellar beat, decreasing motility (Turner, 2006). Furthermore, there are complexes signaling pathways that must work in a proper manner to support sperm motility. In other words, any problem in those pathways could be responsible for poor motility. The number of sperm cells

per ejaculate is another trait used to evaluate que semen and it is directly linked to the number of insemination doses produced by an AI station, influencing its profitability.

Delving deeper into the causes of low semen quality, the studies of genetic relationships among semen traits and the genetic mechanisms controlling these traits appear as important factors. Furthermore, the genetic selection of boars for those traits would be of valuable application. In this context, currently, new traits have been considered in pig breeding programs and, besides the importance of evaluation of semen traits for AI industry economic profitability, genetic selection for semen traits and boar fertility is still not common practice in many breeding companies.

In this thesis, I estimated moderate heritabilities for four semen traits (motility, progressive motility, number of sperm cells per ejaculate and morphological abnormalities) in sire and dam lines. Favorable genetic correlations were found among these traits, such as between motility or progressive motility and morphological abnormalities. In addition, I verified that is possible to genetically select boars for low variation in semen quality and production. These results showed that pig breeding programs should consider the inclusion of semen traits in their selection indexes, which would reduce culling of genetically proven boars in AI stations that do not produce semen in a standard quality or quantity, decreasing costs (Schulze et al., 2014). The high and favorable genetic correlations indicate in special the progressive motility as the best semen trait for assessment of semen quality. Regarding semen quantity, the total number of sperm cells per ejaculate is also an essential trait, considering its genetic independence from the other semen traits.

Genomic era

The sequencing of pig genome and commercial availability of the Porcine Illumina SNP60 BeadChip (Ramos et al., 2009) have marked a new era of the pig breeding industry (Merks et al., 2012), with the application of GWAS and Genomic selection for a wide range and diverse pig traits. New technologies and statistical methods are being developed to improve the use of genomic tools. For example, besides the decreasing costs to genotype animals over the years, the availability of a high number of animals with both genotypes and phenotypes may be an issue for a powerful GWAS. In this circumstance, the single-step GWAS (Wang et al., 2012) may perform better than other GWAS methods. After verifying that selection of boars for semen quality and quantity traits could result in reasonable genetic

progress in Chapter 2, I performed a weighted single-step GWAS to find QTL regions and SNP markers associated with semen traits in Chapter 3. The weighted method is useful because different weights can be considered for SNPs depending on their association with the studied trait and important genomic regions can be pinpointed more accurately. Important candidate genes reported in literature involved in controlling the production process or maturation of spermatozoa were found in QTL regions that explained major proportions (up to 10.8%) of genetic variance for evaluated traits. Although the causal mutations are not pinpointed in GWAS, these studies are useful to increase the elucidation of genes and genetic markers associated to target traits. In addition, they contribute to the amount of studies in the field, helping researchers to compare and have basis for their own studies, especially when they are scarce, as for semen traits.

Paternal side of reproductive traits

With the advances in pig breeding programs, since 1990s reproductive traits, mainly litter size at birth, started to be considered in selection goals (Merks et al., 2012). Especially for dam lines, these traits are commonly treated and selected as traits of the sow, although the service sire may impact their outcome (Van der Lende et al., 1999; Serenius et al., 2003; Hamann et al., 2004; Su et al., 2007; Wolf and Wolfová et al., 2012a,b). A wide range of authors studied the influence of semen traits on boar fertility (Gadea et al., 2004; Gadea, 2005; Ruiz-Sánchez et al., 2006; Vyt et al., 2008; Novak et al., 2010; Broekhuijse et al., 2012a,b) especially for total number born (TNB) and farrowing rate (FR). According to Van der Lende et al. (1999), the fertilization and prenatal survival rates, which control the litter size, might be influenced by the service sire due to differences in the capacity of fertilization, which can be related to sperm quality and/or the boar genetic contribution to viability of the embryo. Regarding the sperm quality, in healthy boars with semen quality that meets AI industry standards, and due to the high number of sperm cells presented in ejaculate doses, the semen analysis are not the best predictor of boar fertility (Ruiz-Sánchez et al., 2006) because variation in TNB and FR can still arise. Other factors that may influence boar fertility are the optimal oestrus detection, the method used for evaluation of semen traits (CASA or technician in a microscope), time of insemination, interactions of female genital tract with the sperm cells and interactions of the semen with the oocyte (Vyt et al., 2008). Finally, the genetic contribution of the service sire might also be an important factor. In Chapter 4, I

showed that there is genetic variation due to service sire and ejaculate effects on GL, TNB and log-transformed number of stillborn piglets (lnSB). Therefore, the genetic paternal side is also important in the control of reproductive traits and should be included in their genetic evaluation models.

Conclusion

In this thesis, the link of traditional and basic genetic evaluation of semen traits in pigs with genomic information resulted in a complete study of the importance and role of the boar in pig breeding industry. It was shown that selection for improved semen quality and production is feasible and the genomic regions and genes found associated with semen traits confirm that the genetic paternal side of these traits should be considered in breeding objectives. Further, the importance of the paternal effect was supported by the variation found for genetic service sire effect and ejaculate on female reproductive traits. This thesis brings up the importance of the male reproduction field in swine and presents innovative scientific information that will increase the still scarce knowledge about genetic selection and genomic architecture of boar semen quality and fertility traits.

References

ALM, K.; PELTONIEMI, O. A.; KOSKINEN E.; ANDERSSON, M. Porcine field fertility with two different insemination doses and the effect of sperm morphology. **Reproduction in Domestic Animals**, v.41, p.210-213. 2006.

BRACKETT, B. G. Reprodução em Mamíferos do Sexo Masculino. **In: REECE, W. O. Dukes Fisiologia dos Animais Domésticos**, Rio de Janeiro: Guanabara Koogan, p.638, 2006.

BROEKHUIJSE, M. L. W. J.; ŠOŠTARIĆ E.; FEITSMA, H.; GADELLA, B. M. Application of computer-assisted semen analysis to explain variations in pig fertility. **Journal of Animal Science,** v. 90(3), p.779–789, 2012.

BROEKHUIJSE, M. L. W. J.; ŠOŠTARIĆ E.; FEITSMA, H.; GADELLA, B. M. The value of microscopic semen motility assessment at collection for a commercial artificial insemination centre, a retrospective study on factors explaining variation in pig fertility. **Theriogenology,** v.77, p.1466-1479, 2012b.

DINIZ, D. B.; LOPES, M. S.; BROEKHUIJSE, M. L. W. J.; LOPES, P. S.; HARLIZIUS, B. GUIMARÃES, S. E. F.; DUIJVESTEIJN, N.; KNOL, E. F.; SILVA, F. F. A genome-wide association study reveals a novel candidate gene for sperm motility in pigs. **Animal Reproduction Science**, v. 151. P. 201-207, 2014.

GADEA J. Sperm factors related to in vitro and in vivo porcine fertility. **Theriogenology**, v. 63, p.431-444, 2005.

GADEA, J.; SELLÉS, E.; MARCO, M. The predictive value of porcine seminal parameters on fertility outcome under commercial conditions. **Reproduction in Domestic Animals,** v.39, p.303-308, 2004.

HENDERSON, C.R. Selection index and expected genetic advance. **Statistical Genetics and Plant Breeding**, National Academy of Science-National Research Council, Washington, DC, 982, p.141-163, 1963.

HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, p.423-447, 1975.

HAMANN, H.; STEINHEUER, R.; DISTL, O. Estimation of genetic parameters for litter size as a sow and boar trait in German herdbook Landrace and Pietrain swine. **Livestock Production Science**, v. 85, p.201–207, 2004.

HERMESCH, S.; LUXFORD, B. G.; GRASER, H. -U. Genetic parameters for piglet mortality, within litter variation of birth weight, litter size and litter birth weight. **Proc. Association for the Advancement of Animal Breeding and Genetics.** p 211-214. 2001.

KNOX, R. V. Artificial insemination in pigs today. **Theriogenology,** v. 85, p.83–93, 2016.

LOPES, M. S. **Genomic selection for improved crossbred performance.** PhD thesis. Wageningen: Wageningen University & Research, 2016.

MAES, D.; LÓPEZ RODRÍGUEZ, A.; RIJSSELAERE, T.; VYT, P.; VAN SOOM, A. Artificial insemination in pigs. In Artificial insemination in farm animals ed. M. Manafi. **InTech**, Rijeka, Croatia p.79–94, 2011.

MERKS, J. W. One century of genetic changes in pigs and the future needs. **BSAS occasional publication**, p.8-19, 2000.

MERKS, J.; MATHUR, P.; KNOL, E. New phenotypes for new breeding goals in pigs. **Animal,** v.6, p.535-543, 2012.

NOVAK, S.; RUIZ-SANCHEZ, A.; DIXON, W. T.; FOXCROFT, G. R.; DYCK, M. K. Seminal plasma proteins as potential markers of relative fertility in boars. **Journal of Andrology**, v.31. p.188–200, 2010.

RAMOS, A. M.; CROOIJMANS, R. P. M. A.; 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. L.; SCHNABEL, R. D.; VAN TASSELL, C. P.; TAYLOR, J. F.; WIEDMANN, R. T.; SCHOOK, L. B.; GROENEN, M. A. M. Design of a high density snp genotyping assay in the pig using snps identified and characterized by next generation sequencing technology. **PLoS One**, v. 4, n. 8, p. e6524, 2009.

ROBINSON, J. A. B.; BUHR, M. M. Impact of genetic selection on management of boar replacement. **Theriogenology,** v. 63(2), p.668–678, 2005.

RODRIGUEZ, A. L.; SOOM, A. V.;. ARSENAKIS, I.; MAES, D. Boar management and semen handling factors affect the quality of boar extended semen. **Porcine Health Management**, v.3, p.15, 2017.

RUIZ-SÁNCHEZ, A. L.; O'DONOGHUE, R.; NOVAK, S.; DYCK, M. K.; COSGROVE, J. R.; DIXON, W. T.; FOXCROFT, G. R. The predictive value of routine semen evaluation and IVF technology for determining relative boar fertility. **Theriogenology,** v.66, p.736–748, 2006.

SCHULZE, M.; BUDER, S.; RÜDIGER, K.; BEYERBACH, M.; WABERSKI, D. Influences on semen traits used for selection of young AI boars. **Animal Reproduction Science**, v.148, p.164-170, 2014.

SERENIUS, T.; SEVÓN-AIMONEN, M. -L.; MANTYSAARI, E. A. Effect of service sire and validity of repeatability model in litter size and farrowing interval of Finnish Landrace and Large White populations. **Livestock Production Science**, v.81, p.213–222, 2003.

SU G.; LUND, M. S.; SORENSEN, D. Selection for litter size at day five to improve litter size at weaning and piglet survival rate. **Journal of Animal Science**, v.85, p.1385–1392, 2007.

TURNER, R. M. Moving to the beat: a review of mammalian sperm motility regulation. **Reproduction, Fertility and Development**, v.18. p.25-38, 2006.

VAN DER LENDE, T.; WILLEMSEN, M. H. A.; VAN ARENDONK, J. A. M.; VAN HAANDEL, E. B. P. G. Genetic analysis of the service sire effect on litter size in swine. **Livestock Production Science**, v.58. p.91–94, 1999.

VYT, P.; MAES, D.; QUINTEN, C.; RIJSSELAERE, T.; DELEY, W.; AERTS, M.; KRUIF, A. DE; SOOM, A. VAN. Detailed motility examination of porcine semen and its predictive value towards reproductive performance in sows. **Vlaams Diergeneeskundig Tijdschrift**, v.77, p.291–298, 2008.

WANG, H.; MISZTAL, I.; AGUILAR, I.; LEGARRA, A.; MUIR, W. M. Genome-wide association mapping including phenotypes from relatives without genotypes, **Genetics Research**, v. 94 p.73-83, 2012.

WOLF J.; WOLFOVÁ, M. Effect of service sire on litter size traits in Czech Large White and Landrace pigs. **Czech Journal of Animal Science**, v.57, p.220-230, 2012a.

WOLF J.; WOLFOVÁ, M. Genetic parameters including the service sire effect for the sow traits stillbirth and piglet losses in Czech Large White and Landrace. **Czech Journal of Animal Science**, v.57. p.402-409, 2012b.