THE UNIVERSITY *of* EDINBURGH

Genetics of Litter Size and Prenatal Survival in pigs

Silvia Clara Hernández Velasco

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Edinburgh

October 2011

This thesis is dedicated to my family,
for their unconditional support.

Table of Contents

List of Figures

List of Tables

Preface

Caminante

Caminante, son tus huellas
el camino, y nada más;
caminante, no hay camino,
se hace camino al andar.
Al andar se hace camino,
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar.
Caminante, no hay camino,
sino estelas en la mar.

Wanderer, your footsteps are
the road, and nothing more;
wanderer, there is no road,
the road is made by walking.
By walking one makes the road,
and upon glancing behind
one sees the path
that never will be trod again.
Wanderer, there is no road
Only wakes upon the sea.

Antonio Machado

"It makes all the difference whether one sees darkness through the light or brightness through the shadows."

David Lindsay

Declaration of Originality

I hereby declare that the work presented in this thesis and the thesis itself have been composed and originated by myself, unless otherwise specified.

Silvia Clara Hernández Velasco

October 2011

Acknowledgements

Firstly I would like to thank my supervisors, Alan Archibald, Cheryl Ashworth and Chris Haley, for their support, advice, and guidance throughout my PhD, thank you for believing in me. Thanks for having the doors always open for my continuous questions. I tried to make the most of your expertises in the different fields involved in this project, and bring them together, I think I have managed. Thank you for always getting excited with the results and bringing back my motivation after a long week of work.

It was a great experience and an excellent environment at the Roslin Institute, especially in the old building; it was the best you can have. This has made easier and a pleasure the difficult path. For this reason, I would like to thank everybody that have made it possible.

In particular, I would like to thank both laboratories where this work was done. From the Archibald group, I would like to thank Alison, Susan, and Jennifer for their unconditional help and special thanks to Heather for teaching me all about genotyping and its complexity, and for all the microsatellites she has genotyped and that have been used in this study. From the Ashworth group, I would like to thank the only and essential member, Charis, thank you for all the help setting all the experiments and for the continual troubleshooting. The numerous tea's including a talk in your office, in company of Claire and Lorna, were essential to keep me going.

I would like to thank Andy for his help with the genotyping analysis in unix and for making easier the verification of the results. Part of the tissues used in this analysis was collected in the Roslin Institute farm, thanks to the collaboration of the stuff from the farm. I would also like to thank Bob for all the help and for introducing me in the world of bioimaging and Image J. I am thankful to Steve and Ozzie, for all the guidance and assistance with the statistical analysis and with the genome-wide analysis.

I would specially like to thank Tahar, for his guidance and encouragement, especially at the beginning of the project, thank you for your positivity and for the help with presentations and abstracts. I learnt a lot from you, and I think now I can tell a story with my work.

One of the most difficult parts of this PhD was the collaboration with France to collect tissues. In this respect, I would like to thank Jean-Pierre, Marie-Pierre, and Yvon for providing me with the tissues and their time, and to the people in the INRA farm in Le Magneraud for the help in the lab despite the difficulties in communication due to the language. I would like to thank Tim for the help with the license preparation for importing the tissues. I have to thank specially Brian and Gerry for driving all the way down to collect my samples and making sure they arrived safely to the Institute, thank you very much for making the effort and sorry for all the problems we have had with the preparatives. I have learnt a lot from this experience, mainly, if you want something to be done and well, better to do it yourself.

I would like to thanks Dr. Larry W. Fisher of the National Institute of Health for providing the rabbit polyclonal LF-166 and LF-124 to human recombinant osteopontin, used in the immunohistochemistry analysis.

Additionally, I want to acknowledge all the people I have met at the Roslin Institute during this period, and that in one way or another have made a difference in this period. Thanks to Richard and Anna for been such good officemates and for those random conversations about the future; thanks to Catalina, Chiara, and Stephanie, for having always been there, reminding me that I should have fun and get some distraction; thanks to Valentina for all the help during the hard work of the last year, despite her own heavy work responsibilities; and to the numerous people that have past through (and/or still are in) the Institute in these years, Bruno, Claudia, Cecile, Andreia, Ciupi, Giulietta, Stefania, Cristina, Ricardo, and Ronan.

I would like to mention all the people I met during this period in Edinburgh outside the Institute. In particular, I would like to thank Pilar for her help at the very beginning of my stay here and for having always been there; also Aida, for being such a good flatmate during the first 6 months and a really good friend after, even if you abandoned me, see you soon; and Olga, my flatmate during the last period in Edinburgh, for her continuous support. I would also like to thank Eva, Julia, and Natalia, thank you for crossing my live.

I would also like to thank my friends from my two homes, San Sebastian and Madrid, for their constant support and encouraging, for keeping me in your minds and receiving me with open arms after months without seeing each other, like if I have been there all this time. In particular, I would like to thank Nacho for all the advices in picture analysis and help with Image J. Thank you all for being there all this time. Now I am coming back, get ready.

Finally, but most importantly, I would like to thank my family for their constant encouragement and support in all forms. Thank you for asking questions and trying to understand what I was working on. I could not have done it without you all. Thank you to my sister, for all the help with the technical difficulties concerning the writing stage, despite her own heavy family and work commitments. Sorry for not being there to help you. Thank you to my mum for coping with all the stress that I brought her way during the hard times. I would like to thank my nieces that without actually doing anything and without realising, gave me a reason to keep going. I am now getting back to normality, and we will be able to do things together once again, trying to recover all this time, I am really looking forward to it.

Globally, it was a great experience despite the challenges of doing a PhD far from home and in another language, but I knew where I was getting in. I have lived in a wonderful and beautiful city, Edinburgh, which I will miss despite the weather, and I had the best company during all the way to achieve my goal. I wouldn't change a thing.

Thank you.

List of publications

Hernandez, S.C., Hogg, C.O., Billon, Y., Bidanel, J.P., Haley, C.S., Archibald, A.L. and Ashworth, C.J. ***Localisation and expression of Secreted Phosphoprotein 1 in uterine and placental tissues from Large White and Meishan pigs.*** Proceedings for the Annual Conference of the Society for Reproduction and Fertility, Brighton 2011. Abstract O24

Hernandez, S.C., Finlayson, H.A., Hogg, C.O., Ashworth, C.J., Haley, C.S. and Archibald, A.L. ***Genetics of litter size and embryo survival in pigs.*** Proceedings for the 32nd Conference for the International Society for Animal Genetics, Edinburgh 2010. Abstract P4028 pp 98

Hernandez, S.C., Hogg, C.O., Khazanchi, N., Haley, C.S., Archibald, A.L. and Ashworth, C.J. ***Localization and expression of SPP1 in porcine uterine and placental tissues.*** Proceedings for the Annual Conference of the Society for Reproduction and Fertility, Nottingham 2010. Abstract P78 p83

Hernandez, S.C., Finlayson, H.A., Ashworth, C.J., Haley, C.S. and Archibald, A.L. ***Mapping Quantitative Trait Loci for reproduction in pigs*** Society of Reproduction and Fertility Supplement, 2009; Vol. 66. Control of Pig Reproduction VIII, Nottingham University Press. pp 117-118

Hernandez, S.C., Finlayson, H.A., Ashworth, C.J., Haley, C.S. and Archibald, A.L. ***Mapping Quantitative Trait Loci for Embryo survival and Litter size in pigs.*** Proceedings for the 31st Conference for the International Society for Animal Genetics, Amsterdam 2008 Poster 2016

Abstract

Female reproductive performance is a critical component of sustainable pig production systems. There is abundant evidence of genetic variation in these traits among pig breeds. The aims of this study were to identify quantitative trait loci (QTL) affecting reproductive traits and to identify and characterise positional candidate gene(s) underlying the QTL. A Large White - Meishan F_2 population was scanned for QTL with effects on reproductive traits. This analysis revealed 13 putative QTLs on seven different chromosomes with effects on five different traits: ovulation rate (OR), teat number (TN), prenatal survival (PS), total born alive (TBA) and litter size (LS). QTL for PS and LS on chromosome 8 were fine mapped and Secreted Phosphoprotein 1 (*SPP1*) confirmed as a candidate gene. A genome-wide association study was performed on a diverse population of different breeds and crosses lines, for reproductive traits including LS, TBA, number of stillborn piglets, and number of mummified piglets. Fourteen SNPs were found significantly associated with reproductive traits.

The functional study of SPP1 examined the hypothesis that differences in foetal growth may be associated with the effectiveness of conceptus attachment, as measured by SPP1 expression. Patterns of SPP1 mRNA and protein expression in placental and uterine tissues supplying the smallest and a normal-sized foetus from the same uterus were examined in Large White-Landrace (LW-LR), Large White (LW) and Meishan (MS) females 40 and 45 of pregnancy. The smallest LW-LR foetuses tended to have a higher level of SPP1 mRNA in endometrium tissue compared to the normal-sized foetuses. However, placenta expression was higher in the normal-sized foetuses compared to the smallest ones. SPP1 protein levels in normal sized foetuses were significantly higher than in the smallest litter mates for all the tissues. Significantly higher levels of SPP1 mRNA and protein were found in MS compared to LW. In both breeds, significant differences between sizes were found in some tissues, with similar expression patterns in respect to size, for both mRNA and protein in endometrial tissues when compared to contemporary LW. In placenta, the direction of the expression differed between breeds, with a higher expression of mRNA and protein in the normal-sized MS foetuses and in the smallest

sized LW foetuses. The comparison of SPP1 expression between different foetal sizes and different breeds revealed associations between breed, foetal size, and SPP1 protein, factors implicated in PS and LS. These results together with the genetic evidence indicate that the potential role of SPP1 in placental and foetal development merits further investigation.

List of Abbreviations

AG1	Age group 1
AG2	Age group 2
AP	Age at puberty
AREG	Amphiregulin
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
CI	Chloroform:Isoamylalcohol
C.I.	Confidence interval
CL	*Corpora lutea*
cM	Centi Morgans
cpm	Count per minute
Ct	Cycle threshold
CV	Co-efficient of variance
dDW	Doble deionised water
DEPC	Diethylpyrocarbonate
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide 5'-triphosphate
DPX	Distyrene, plasticiser and Xylene
dsDNA	Double-stranded DNA
DTT	dithiothretiol
E	Amplification efficiency
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
End	Endometrium
EPOR	erythropoietin receptor
ES	Embryo survival
ESR	oestrogen receptor
EtOH	Ethanol
FFPE	Formalin-fixed, Paraffin Embedded
FPR	Ratio of foetal weight and placental weight
FSHβ	follicle-stimulating hormone β
g	Gravitational
GAPDH	Glyderaldehyde 3-phosphate dehydrogenase

GE	Endometrial glandular epithelium
GF	Glass microfiber binder
GL	Gestation length
GNRHR	Gonadotropin releasing hormone receptor gene
GWAS	Genome-wide association study
H&E	Haematoxylin and Eosin
HCl	Hydrochloride
Hi Di	Highly deionized
HSA	Human H*omo sapiens* chromosome
IBS	Identicla-by-state
IgG	Immunoglobulin
IHC	Immunohistochemistry
IPTG	Isopropyl-beta-D-thiogalactopyranoside
ISH	*in situ* Hybridisation
LacZ	β-galactosidase gene
LB	Lysogeny broth
LE	Endometrial luminal epithelium
LL	Lactation length
LOD	Logarithm (base 10) of odds
LR	Landrace
LS	Litter size
LW	Large White
M	Gene-stability measurement
MAF	Minor allele frequency
MARC	Meat Animal research center
MAS	Marker-assisted selection
Mbp	Mega base pairs
MBq	Mega Becquerel
mRNA	Messenger ribonucleic acid
MS	Meishan
NaAc	Sodium Acetate
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NF	Normalisation factor
NFF	Number of fully formed piglets
NMUM	Number of mummified piglets
Nor	Normal
NSB	Number of stillborn piglets
NTC	no template control

NVE	Number of viable embryos
NWEAN	Number of piglets weaned
O.C.T.	Optimal Cutting Temperature
OD	Optical density
OR	Ovulation rate
PBS	Phosphate-buffered saline
PCI	Phenol:Chloroform:Isoamylalcohol
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
Plac	Placenta
PRLR	prolactin receptor
PS	Prenatal survival
PTGS2	prostaglandin-endoperoxido synthase 2
Q	quantity
qPCR	Quantitative PCR
QTL	Quantitative trait loci
R^2	Correlation coefficient
rATP	Riboadenosine 5'-triphophate
RBP4	retinol-binding protein 4
rCTP	Ribocytidine 5'-triphosphate
RFLP	Restriction fragment length polymorphism
RGD	Arginine-Glycine-Aspartic
rGTP	Riboguanosine 5'-triphosphate
RIN	RNA integrity number
RNA	Ribonucleic acid
RNase	ribonuclease
RNasin	RNase inhibitor
rNTP	Ribonucleotide 5'-triphosphate
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcriptase
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
rUTP	Ribouridine 5'-triphophate
SD	Standard deviation
SDCt	Standard deviation of cycle threshold
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean
Sm	Smallest
SNP	Single nucleotide polymorphism

SOC	Super optimal broth
SPP1	Secreted phosphoprotein 1
SSC	*Sus scrofa* chromosome
TAE	Tris base, glacial acetic acid and EDTA
TBA	Total number of piglets born alive
TBD	Total number of piglets born dead
TCA	Trichloroacetic acid
TE	Tris and EDTA buffer
Tm	Annealing temperature
TN	Teat number
TNB	Total number of piglets born
Tr	trophoblast
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
U.V.	Ultra violet
UC	Uterine capacity
UCSC	Univerity of California, Santa Cruz
UHO	Unilateral hysterectomy and ovariectomy
Uter	Whole utero-placental unit
USDA	United States Department of Agriculture
V	Pair-wise variation
VEGF	Vascular endothelial growth factor
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

Chapter 1
General Introduction

1.1. Pig production and the implications of reproduction

Reproduction, especially female reproductive performance, is an important component in livestock production. In pig production, this factor is of special interest, as pork is an important and increasing source of food in the world (Orr Jr & Shen, 2006). Together with the increase in pig production for the market, there is an increase in the cost of this production in the pork industry due to diverse reasons, such as the effects of environment changes, the cost of feed for these animals and the demands of the consumer for a quality product (BPEX annual technical report 2010-2011; http://www.ukagriculture.com/). During 2010 an improvement in mortality was displayed, however the bad weather conditions in UK restricted other improvements in performance. Thus, despite the increase in number of pigs reared per sow per year and in weight of pigs produced from 2006 (21.5 pigs and 98.2 Kg) to 2010 (22.1 pigs and 105.2 Kg), other factors had reduced the profit in pig production (BPEX annual technical report 2010-2011). The number of heads in breeding herds in the UK was 424,500 in 2010, which produced a total of 775,000 tonnes of meat, 75,000 tonnes less than in 2006 when the number of head was 458,000 (BPEX annual technical report 2010-2011). Despite this increase in the production during the last few years, the total number of heads used for production had been reduced significantly during the last 20 years (DEFRA). Ashworth *et al.* (2004) estimated a sustainable £ 1m benefit for the UK pig industry with an increase of one piglet per sow per year. Thus, an increase in the number of piglets a sow can farrow could translate in a reduction of the sows per herd and as a consequence the profit increases in pork production, maintaining the quality of the product and the permanent cost constant.

For this reason, in the last decade, selection for improved prolificacy has been performed in different countries with a consequent moderate increase of litter size (LS) at birth (Bee, 2007). However, this increase in number of piglets at birth has led to an increased within-litter variation in birth weight, as well as a decrease in the birth weight per piglet. These effects have been associated with greater preweaning

mortality, slower growth rates, and decreased pork quality (Quiniou *et al.*, 2002; Herpin *et al.*, 2002; Foxcroft *et al.*, 2007). As a consequence of this decrease in birth weight, the competence of the piglets decreases, not only for growth but also for survival (Foxcroft *et al.*, 2007). This translates into a reduction of litter size or poor performance, both causing a cost increase or reduction in productivity. Thus, LS and developmental competence need to go together as factors for selection for improvements in production.

There is prenatal programming of postnatal development in the pig, meaning that the size of the piglet is determined during pregnancy (Foxcroft *et al.*, 2009). Thus, factors affecting the development of the embryo and foetus need to be taken into account. This include maternal and environmental effects, as well as uterine and conceptus factors. The understanding of these factors offers the opportunity for an effective increase in LS at term of pregnancy and for increased productivity as the pigs grow.

1.2. Pigs in agriculture and as an animal model in medicine

Pigs are multiparous animals with a generation interval of 1 year, a gestation period of 114 days, and large litters of ten or more piglets. The ideal gilt reaches puberty at approximately 6 month of age, farrows a large number of progeny per litter, promptly returns to oestrus, and can be successfully breed for many parities (Rohrer *et al.*, 1999). However, both prenatal and postnatal losses limit opportunities for greater economic profitability of the swine industry. Therefore, by increasing the number of potential piglets per litter that an individual sow farrows, the size of the stock of females can be reduced producing a more efficient production system, with increased outputs and reduced overhead costs and environmental footprints. As well as high levels of reproductive success and survivability, the pork industry requires efficient growth rates, reduced feed intake, better carcass merit, and meat quality due to their economic value for meat production (Rothschild, 2003).

Genetic differences in reproduction have been observed both among breeds and lines. Hence, while progress can be made using conventional selection, marker-assisted selection (MAS) offers an opportunity to improve selection programmes for production and reproductive traits (Spotter & Distl, 2006). A primary focus of the animal genetics field is the elucidation of genes influencing diverse phenotypes of both agricultural and biomedical relevance (Meyers *et al.*, 2005).

Furthermore, as omnivores, with a cardiovascular and gastrointestinal physiology similar to humans in structure and function, pigs have often been used as a biomedical research model to help understand human physiology, behaviour, and disease. Pigs are sufficiently similar to humans for them to represent a significant future source of organs for transplantation (xenotransplantation) (Cooper *et al.*, 2008).

Finally, advances in genomics are creating opportunities to understand the genetic control of complex traits including reproductive performance. Therefore, a genome sequence for the species of interest is considered essential for modern biological research. The pig (*Sus scrofa*) genome is formed of 18 autosomal pairs plus an X/Y pair of sex chromosomes, and it has an estimated size of 2.7×10^9 bp (Humphray *et al.*, 2007), similar to that of the human genome, which is 3.2×10^9 bp (Venter *et al.*, 2001; McPherson *et al.*, 2001). The Swine Genome Sequencing Consortium is close to publishing a draft pig genome sequence (Archibald *et al.*, 2010).

1.3. Reproduction

In males, reproductive performance traits include testis size, semen volume, sperm concentration of the ejaculate, sperm quality, and libido or breeding aggressiveness. Reproductive traits in females include age at puberty (AP), oestrous cycles and oestrus expression, LS, weaning to oestrus interval and farrowing interval. The component traits of LS are ovulation rate (OR), fertilisation rate, prenatal survival (PS), gestation length (GL), uterine capacity (UC), embryo survival (ES), number of born alive piglets (TBA), number of stillborn piglets (NSB), number of mummified

piglets (NMUM), and preweaning losses. As the capacity to feed multiple offspring is critical to ensuring that the piglets survive and grow, number of teats (TN) is also a trait that contributes to reproductive performance.

The pig is a spontaneously ovulating mammal with an oestrous cycle of 21 days. Puberty is reached at age of 190 days, with a first puberty oestrus. However, the first mating will occur usually one or two oestrous cycles after puberty, due to the increase in number of ova with the number of cycles and a consequent increase in LS when mated at this stage.

1.3.1. Pregnancy in pigs

Early pregnancy in pigs is a complex process influenced by the overlapping events of conceptus (embryo/foetus and associated extra embryonic membranes) elongation, endometrial remodelling for implantation and pregnancy recognition signalling. Establishment and maintenance of pregnancy result from signalling by the conceptus oestrogen and requires progesterone produced by the *corpora lutea* (CL) (Spencer *et al.*, 2004). Progesterone increases the expression of various uterine secretion proteins to support early development events (Spencer *et al.*, 2004).

After copulation or insemination, the sperm travel to meet the oocytes for fertilisation. After fertilisation of the ova (day one of pregnancy), the cell divisions start and after five days the blastocyst is formed (Figure 1.1). The blastocyst consists of an inner cell mass which will form the embryo, and an outer layer of cells, or trophoblast (Tr), that will give rise to the placenta. The blastocyst expansion and elongation begins and at the same time the blastocyst moves around the bicarnate uterus (formed by 2 horns) to find an implantation site around day 12 to 22. This period from fertilisation to implantation is defined as the pre-implantation period. Once the blastocyst is implanted, it is defined as an embryo until around day 30, thereafter it is called a foetus. The uterus plays a key role in maternal recognition of pregnancy, embryo elongation, implantation, and support of the developing embryo during this period of early pregnancy (Vallet *et al.*, 2002b). The uterus, composed by

the luminal epithelium (LE), glandular epithelium (GE) and the myometrium, secretes a series of nutrients and other substances into the lumen that are collectively known as histotroph.

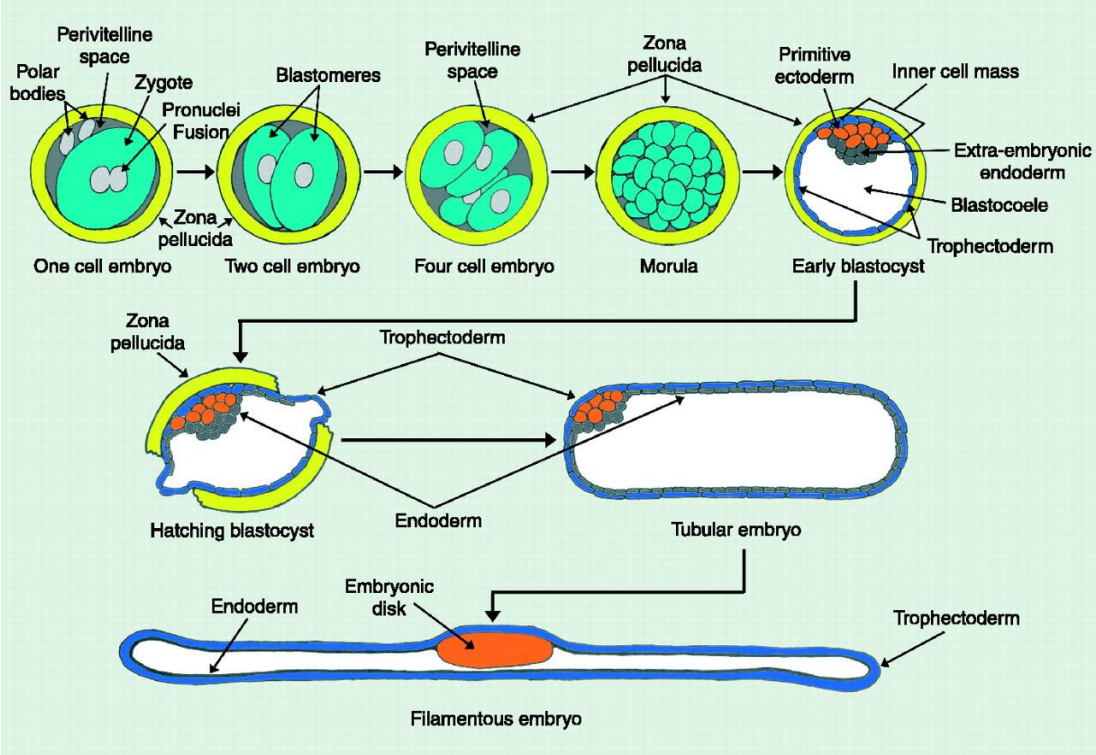

Figure 1.1 Early pregnancy events in domestic animals. Fertilisation occurs in the oviduct, and morula-stage embryos enter the uterus where they develop into spherical blastocysts and hatch from the zona pellucida through the actions of proteases. Thereafter, spherical blastocysts migrate, assume a tubular and then a filamentous form due to rapid elongation of the trophectoderm before initiation of implantation. Implantation involves apposition and transient attachment followed by firm adhesion of trophectoderm to uterine luminal and superficial glandular epithelia. Figure taken from Bazer *et al.* (2009).

1.3.1.1. Elongation of blastocyst and Recognition of pregnancy

Progesterone plays pivotal roles during gestation, including preparation of the endometrium for implantation, maintenance of pregnancy, and uterine quiescence, through a segregation of a complex variety of proteins (Roberts & Bazer, 1988). Proper embryonic development throughout the preimplantation period is a prerequisite for establishing a successful term pregnancy. Elongation of the pig

conceptus is characterised by four morphologically distinct stages of development, which include spherical, ovoid, tubular, and filamentous forms (Anderson, 1978; Geisert *et al.*, 1982; Wilson *et al.*, 2000).

Following the embryonic oestrogen signal for maternal recognition of pregnancy (peri-implantation period), the elongated conceptuses remain free-floating in the uterine lumen until day 13-14 of gestation. Attachment of the embryos to the uterine endometrial luminal epithelium (LE) begins and it is complete 10 days later (Ashworth, 2006). Some embryos migrate from one uterine horn to the other with the objective to achieve even spacing throughout the female reproductive tract (Dantzer & Winther, 2001).

Due to the type of placentation of pigs the contact of the conceptus with the uterus must be of certain area in order to maintain pregnancy. Thus, an inability of the sow exist to maintain a unilateral pregnancy early in gestation (Anderson, 1966) and at least two embryos need to be present in each uterine horn to establish pregnancy (Polge *et al.*, 1966) in order for the oestrogen secreted by the conceptus as a signal of pregnancy to have enough strength to prevent uterine prostaglandin secretion which will induce luteolysis (Ashworth, 2006). This suggests that oestrogen does not diffuse easily throughout the uterine horn (Geisert *et al.*, 1990). A second, more sustained increase of oestrogen secretion is observed between days 15 and 25-30 of pregnancy (Geisert *et al.*, 1990). Inhibition of luteolysis and establishment of pregnancy in pigs require this biphasic pattern of oestrogen secretion that results in prolonged luteal life span and progesterone secretion (Geisert *et al.*, 1990).

There is an endocrine-exocrine theory of pregnancy recognition in pigs. As described previously, it is known that the uterine endometrium secretes progesterone and the conceptus secretes oestrogen, which are antiluteolytic. Thus, this theory states that during a normal cycle in gilts, progesterone is secreted in an endocrine manner into the CL, in order to perform its luteolytic function. However, during pregnancy the secretion of progesterone is exocrine, and it is secreted into the uterine lumen and sequestered in order to prevent its luteolytic function (Bazer, 1992).

1.3.1.2. Attachment of the blastocyst

Implantation consists in the attachment of the blastocyst to the uterus for juxtaposition of embryonic and material circulations, resulting in a successful pregnancy and the establishment of a functional placenta (Denker, 1993; Carson *et al.*, 2000). The period of conceptus attachment and implantation in pigs, as in other species, is a critical time for embryonic survival. Thus, failure of the conceptus to attach properly to the uterine surface may contribute to the high rate of embryonic loss observed in swine and in humans (Ross *et al.*, 2007). In fact, implantation (days 14-19) is one of the periods of greatest embryonic loss in the pig (Pope, 1994). The endometrial LE remains intact throughout pregnancy, and the conceptus trophectoderm attaches to the apical LE surface without invasion of the maternal tissues (Johnson *et al.*, 2003c).

Implantation involves pregnancy-specific remodelling of extracellular matrix (ECM) proteins and adhesion molecules at the conceptus-maternal interface (Carson *et al.*, 2000; White *et al.*, 2006). These two molecules together with cell-surface receptors and growth factor expression change in humans during endometrium transformation to an implantation-receptive state (Carson *et al.*, 2002). These molecules are hypothesised to have direct roles in conceptus attachment, invasion, and placental development. Pregnancy loss due to defects in implantation is a major cause of infertility in humans, and it is of clinical importance to identify the genes, as well as the cellular and the molecular mechanisms that underlie this critical ECM/adhesion molecule-dependent crosstalk between conceptus and uterus (Norwitz *et al.*, 2001; White *et al.*, 2006).

In pigs and sheep, it appears that integrins play a dominant role in these fundamental processes via interactions with ECM molecules and other ligands to transduce cellular signals in uterine epithelial cells and conceptus trophectoderm (Burghardt *et al.*, 2002). Integrins are a family of cation-dependent heterodimeric intrinsic membrane glycoproteins composed of non-covalently linked α and β subunits that bind to various ECM components and cell adhesion molecules (Giancotti & Ruoslahti, 1999). Integrins have been implicated in the porcine implantation cascade

(Garlow *et al.*, 2002). The integrin family of cell adhesion molecules are a major class of receptors for the ECM and participate in cell-cell and cell-substratum interaction. They are expressed in the endometrium, where they exhibit not only a constitutive and cycle-dependent expression but also a hormone-dependent regulation (Lessey, 1995; Spencer *et al.*, 2004). The central role of integrins in the implantation adhesion cascade is a result of their ability to bind ECM and other ligand(s) to mediate adhesion, migration, invasion, and reorganisation. These receptors are present on the plasma membrane as heterodimeric α and β subunits. The integrins present at the time of attachment in the endometrium are α1β1, α4β1, αvβ3. Of these integrins, αvβ3 is the only one that recognises an Arg-Gly-Asp (RGD) peptide sequence. Progesterone increases expression of αvβ6 and α5β1 integrins during the peri-implantation period, which may in part define the "implantation window" in this species (Burghardt *et al.*, 2002; Spencer *et al.*, 2004). Secreted phosphoprotein 1 (SPP1) is a phosphorylated acidic glycoprotein member that contains a RGD sequence and binds to integrins αvβ3 on LE and αvβ6 on trophoblast to induce focal adhesion assembly, adhesion and migration of conceptus trophectoderm cells during implantation (Erikson *et al.*, 2009). SPP1 is also linked to the establishment and maintenance of pregnancy in ewes (Johnson *et al.*, 2003b), and it is a marker for uterine receptivity to implantation in humans (Carson *et al.*, 2002). SPP1 is also found in porcine endometrium with a different spatial pattern to the pattern in sheep (Garlow *et al.*, 2002). This evidence points to SPP1 protein as a key part in the implantation process, not only in pigs but also in sheep and humans, and possibly in other animals.

1.3.1.3. Non-invasive Placentation in pigs

Placentation is the culmination of a complex series of biochemical and structural interactions between the conceptus and the maternal system. Factors affecting early events in placental development are essential for the exchange of nutrients and gases required for survival and growth of trophoblast initially, and subsequently, foetal-placental tissues (Kim *et al.*, 2010). The placenta is a provisional organ, which only emerges during gestation. Despite the diversity of placentation strategies, the initial events of apposition, attachment, and adhesion between maternal uterine LE and

conceptus trophoblast are shared among species (Burghardt *et al.*, 2002). Whereas, other mammals undergo different degrees of LE degeneration and fusion with the trophoblast during placentation, in pigs there is no penetration of the uterine epithelium by the trophoblast. The pig has a non-invasive diffuse folded epitheliochorial placenta, where implantation remains superficial, which means that the maternal blood supply is well separated from the absorptive surface of the chorion and nutrient exchange is limited by placental blood flow.

The placenta provides several critical functions during pregnancy, such as regulating the transport of nutrients, gases and waste, acting as an immunological barrier, and serving as a source of various proteins growth factors and hormones (Regnault *et al.*, 2002). As gestation is short in pigs (114 days), a high rate of exchange is necessary (Vallet & Freking, 2007). The efficiency of these processes affects the development of the pig foetus, and thus influences UC, LS, birth weights, PS and pre-weaning mortality, and postnatal health and survival.

The importance of UC as a contributor to prenatal loss and LS may result from the type of placentation found in this species and placental efficiency (Wilson *et al.*, 1999). Placental efficiency is the ability of the placenta to support foetal growth and development (Vallet *et al.*, 2002a) and it includes the placental vascular density (Biensen *et al.*, 1998; Vonnahme *et al.*, 2002), placental vascular endothelial growth factor (VEGF), which is related to the vascular permeability of the placenta/endothelium (Vonnahme *et al.*, 2001) and the efficiency of foetal erythropoiesis (Vallet *et al.*, 2002a). Placental efficiency in swine is a complicated trait. It is measured as the ratio of foetal weight and placental weight (FPR), which increases with advancing gestation (Vallet & Freking, 2007). Placental efficiency is very low from day 30 to day 40 of gestation, but begins to increase very rapidly by day 50 to accommodate foetal growth, and continues to increase, although at a reduced rate, until term (Wilson & Ford, 2001).

To meet the requirements of rapidly growing pig foetuses, the placenta either increases in attachment surface area over the endometrium to become large or increases in vascular density (Tayade *et al.*, 2007). Without compensatory increases

in physiologically relevant traits for placental function, a decrease in placental weight compromises foetal survival (Mesa *et al.*, 2005). More efficient placentas are usually smaller, more vascularised, and thicker than less efficient placentas (van Rens & van der Lende, 2004). Placental protein secretion responds to variations in uterine space and the requirements for nutrients of each foetus (Vallet & Christenson, 1993). In contrast to placental protein secretion, endometrial protein secretion appeared unaffected by uterine space. Development of the placental vascular architecture is of considerable importance in influencing the exchange of nutrients, oxygen, and carbon dioxide between mother and foetus essential for growth and development of conceptuses, implantation and placentation (Dantzer & Leiser, 1994; Wang & Dey, 2006).

Angiogenesis is a process in which new capillaries develop from the pre-existing vessels. Physiological angiogenesis occurs at the maternal foetal interface. The potential role of placental-derived angiogenic factors in modulating placental efficiency has been investigated. It is possible that breed differences exist in the spatial and temporal production of these angiogenic factors leading to marked differences in placental tissues and thus placental efficiency (Biensen *et al.*, 1999). It would be of great interest to understand the molecular mechanism underlying the differences in placental gene expression between Chinese indigenous and western breeds of pig (Zhou *et al.*, 2009).

1.4. Losses produced during pregnancy

Molecular understanding of embryonic implantation and development is of particular interest for the study of human infertility. Early pregnancy loss in humans, which often occurs due to defects that take place before, during or immediately after implantation, is a worldwide social and economic concern (Wang & Dey, 2006). Estimates of prenatal losses (embryonic and foetal losses) suggest that up to 40% of oocytes shed at ovulation are not represented by piglets at birth. The factors implicated in this loss are numerous and complex, requiring an extensive examination.

Fertilisation is an important factor for production in pigs, but with appropriate reproductive management, and due to its efficiency (95%), it is a component that has little influence on LS (Pope *et al.*, 1990). Thanks to this efficiency of fertilisation, early embryonic mortality and later foetal survival in utero (Geisert & Schmitt, 2002) can be assessed throughout gestation by comparing embryo/foetus numbers to the number of CL. From these fertilised ova around 10% will be degraded due to ovulation of primary oocytes, fertilisation failure, polyspermy, chromosomal aberrations, and abnormal blastocyst development. After this reduction, only the potentially viable embryos are left (Bennett & Leymaster, 1989).

During the elongation and migration in the uterus (between day 5 and 13), a further 10% embryos are lost, leaving around 75% of fertilised ova to start the attachment period at around day 12. These losses are defined as embryonic mortality, which includes all the losses that occur before day 35 of gestation. Most losses at this period are characterised by asynchrony between conceptus signals and uterine receptivity, resulting in defective implantation and/or placentation (Pope *et al.*, 1982) together with variation in blastocyst elongation. The synthesis of oestradiol from day 10 to 12 by the pig blastocyst with a higher amount coming from the most developed embryos, perhaps alters some component(s) of the endometrial histotroph (proteins, growth factors, calcium, and prostaglandins from the endometrium), critical for survival of the lesser developed littermates (Wu *et al.*, 1988; Pope *et al.*, 1990). Pope *et al.* (1990) studied the different factors affecting the variation blastocyst size, finding the less developed ones to be more susceptible to the environment compared to the more mature ones. Furthermore, the length attained by a conceptus after elongation plays a significant role in determining subsequent placental size (Wilson & Ford, 2001).

Losses after this period and up to term are defined as foetal mortality. The foetal development period can be divided in three. First, the early foetal period from day 35 to 40, when there is an increase in placental length. The second period starts shortly after mid gestation from day 55 to 75, when there is an increase in placenta weight. Finally, the third period, prior to farrowing, runs from day 100 to term (van der Lende & van Rens, 2003). Most of the losses that occur during these periods are due

to UC, which becomes limiting, affecting the number of conceptuses dramatically, as well as to placental development (Vallet & Christenson, 1993; Vonnahme *et al.*, 2002; Wesolowski *et al.*, 2004; Foxcroft *et al.*, 2006). Important factors during this period are the ability of the placenta and the uterus to deliver nutrients, the ability of the foetus to efficiently use them, and the limiting UC due to rapid foetal growth from day 21 to 45 (Pope *et al.*, 1990).

To illustrate these losses, in a more numerical way, an example is used here. Imagine the initial number of ova is between 22 and 23. Out of these, only 17 will arrive to implantation, after embryonic losses. Subsequently, after losses during the foetal period, only 13 to 14 foetus will survive to term (Ashworth, 2006).

The importance of UC may result from the type of placentation found in this species (Wilson *et al.*, 1999). The efficiency of placental attachment and the supply of nutrient to the foetus, also appears to be critical to ensure appropriate foetal growth (Ashworth *et al.*, 2001). Thus, exploration of the genetic factors that regulate placental efficiency is an important research area (Zhou *et al.*, 2009). A foetus less than two-thirds of the average weight of foetuses in the same horn, not only had a lower weight than its siblings but also had a smaller placenta with less blood flow. Moreover, birth size affects the long-term health of an individual, and it is critical in determining life expectancy (Zhou *et al.*, 2009).

There are also some losses during the last days of pregnancy. These losses will be seen as fully formed foetus at birth and they are defined as stillborn piglets. This term also includes the piglets that die during farrowing and just after it. Important risk factors are birth weight and within-litter variation in birth weight (Roehe & Kennedy, 1993; van Rens *et al.*, 2005) with the smallest piglets at greatest risk of dying during parturition piglets (Vallet *et al.*, 2002b). NSB can account for 10-15% of total mortality (Vallet *et al.*, 2002b) and, as a result, the viability of piglets in utero and during parturition are significant issues, together with piglet weight at birth. The perinatal survival chances of a piglet mainly depend on its ability to cope with stresses experienced during farrowing and during the first days after birth (Randall, 1978; Arai *et al.*, 2001; Leenhouwers *et al.*, 2002).

1.5. Factors affecting Litter size

LS at birth and TBA piglets have been viewed as the most economically important reproduction traits to the swine industry (Rothschild, 1996). Measuring LS is straightforward and it is highly variable both between and within breeds, and currently varies from approximately 2 to 20 pigs per litter, with means from 9 to 11 (Irgang *et al.*, 1994). However, it must be taken into consideration that it is a composite trait, made up of many traits expressed by the embryo and the dam (Linville *et al.*, 2001). This trait is also influenced by environmental factors.

Although OR sets the limit for potential LS in swine, LS is clearly not limited by OR within the modern swine herds (Cunningham *et al.*, 1979; Lamberson *et al.*, 1991). OR is an estimate of the number of CL in both ovaries and it can be measured by laparoscopy after fertilisation, thus sometimes the measure does not correlate with LS. OR is also greatly influenced by breed type and is open to improvement by genetic selection (Johnson *et al.*, 1999d). For an actual measurement of number of CL, the dissection of the ovary is required. Sometimes the number of CL is lower than the actual number of embryos or piglets. There are several possible explanations for this phenomenon. Firstly, numbers of CL may have been underestimated. Secondly, it is possible that a follicle may release more than one oocyte at the time of ovulation. An alternative explanation is that monozygotic twinning occurs to a greater extent that previously assumed (Ashworth *et al.*, 1998). An increase in OR has achieved a greater number of embryos at day 30 of gestation. However, LS at parturition has not been increased significantly. These findings suggest that UC must also be improved to increase LS (Christenson *et al.*, 1987).

UC has been defined as the maximum number of foetuses that can be carried successfully to term, when the number of potentially viable foetuses is not limiting (Christenson *et al.*, 1987). However, Wilson *et al.* (1999) suggested that UC in pigs should be defined more correctly as the total amount of placental mass or surface area that a dam can support to term. In this definition the physical space as well as the function is included. Therefore, UC is a major component contributing to LS in

pigs with an important effect on the survival of the foetuses during pregnancy (Christenson *et al.*, 1987). Several experiments have implicated conceptus development during early pregnancy (Youngs *et al.*, 1994), placental efficiency (Biensen *et al.*, 1998; Wilson *et al.*, 1999), and foetal erythropoeisis as possible factors influencing UC of pigs (Vallet *et al.*, 2002a). The incompatibility of increased numbers of conceptuses surviving to the post implantation period, in the absence of increased UC, offers a biological explanation for increased variability in birth weight and postnatal growth performances reported in later parity sows (Foxcroft *et al.*, 2006). However, UC is a difficult trait to measure and together with placental size has a significant influence on foetus weight (Vianna *et al.*, 2004).

A technique suggested to measure the UC was the unilateral hysterectomy and ovariectomy (UHO), which consisted of removing one ovary and horn of the uterus. A UHO female ovulates similar numbers of ova from the remaining ovary as would an intact female, but has approximately one-half the uterine space (Wilson *et al.*, 1999) as a compensatory mechanism (Christenson *et al.*, 1987). Therefore, as LS was independent of the OR in these gilts, the UC for one uterine horn determined LS (Christenson *et al.*, 1987).

However, piglets vitality is also an important factor, since survival is reduced in low birth weight piglets, which display poor thermoregulatory abilities and are slow to acquire colostrum (Baxter *et al.*, 2008). Published studies indicate that a considerable amount of the variation in growth performance after birth may be largely determined, and essentially pre-programmed, during foetal development in the uterus (Foxcroft *et al.*, 2006). The within-litter variation in weight during the pregnancy can be reflected in the variation of weight at birth of the different piglets in a litter. These differences are the cause of the low viability of low birth weight piglets that have to compete with their littermates which are stronger, and this variation is positively related to pre-weaning mortality (Ashworth, 2006).

Some other obstacles to genetic improvement of LS may be maternal effects, intrauterine environment, milk production, and mothering ability of the dam, which may affect her offspring's reproductive performance. Thus, the physical conditions

of the sow at farrowing play an important role in postnatal survival (Vianna *et al.*, 2004; Jonas *et al.*, 2008). For example, TN plays a significant role when many piglets are born but it is not a determinant of LS (Bennett & Leymaster, 1989; Mesa *et al.*, 2003). The within-litter variation in the birth weight and piglets weight gain during suckling has been also related with maternal genetic variation with an heritability similar to that of LS (Damgaard *et al.*, 2003). During the first weeks of life of the piglets there is a great risk of been crushed by the mother or die due to starvation (Weary *et al.*, 1996). Thus, both the developmental competence of the pigs born, as well as the size of the litter, and the condition of the sow at farrowing needs critical consideration (Hellbrugge *et al.*, 2008). Therefore, the welfare of the sow and the litter is an important factor to take in account when increasing LS (Rutherford *et al.*, 2011)

1.6. Meishan characteristics: Ovulation rate, uterine capacity, and placental efficiency compared to other breeds.

The Chinese Meishan (MS), a member of the Taihu group of breeds, is one of the most prolific pig breeds known, farrowing between three to five more piglets per litter than the European commercial breeds (Haley *et al.*, 1995). Moreover, MS females reach puberty around 3 months earlier than European breeds and are considerably smaller at puberty. However, the MS is not commercially viable in Europe, due to its poor growth rate and high carcass fat content. An understanding of the causes and regulation of this prolificacy would not only be of value from a scientific point of view, but might also help with understanding the key mechanisms in other species, and indicate ways in which the prolificacy of other pig breeds could be enhanced.

In comparisons of the OR between MS and Large White (LW), a European commercial breed, a difference was found between the studies in France and the UK. The French studies found either no difference between the breeds or a greater OR in the LW females (Bolet *et al.*, 1986; Bazer *et al.*, 1988). In contrast, in the UK

population, the MS females showed a greater OR than LW ones, both as gilts and sows (Haley & Lee, 1990; Ashworth *et al.*, 1992). OR was found to be similar in gilts at comparable number of oestrous cycles, but in older sows, MS have a higher OR than LW pigs and the difference between breeds seems to increase as the sows get older (Christenson *et al.*, 1987; Bennett & Leymaster, 1989; Haley & Lee, 1993). Similar results were found in studies in the US using American breeds and MS (Christenson *et al.*, 1993).

Several factors may account for the difference between the studies in the UK and France. First, the MS population in France was originally established from two gilts and one boar, a very limited sample of animals (Bazer *et al.*, 1988; Bidanel *et al.*, 1989), compared with 21 females and 11 boars in the importation into the UK (Haley *et al.*, 1992). Second, as the animals imported into the UK and France were obtained from different stock farms, they may be derived from different Chinese subpopulations of MS pigs. Finally, there may be effects of differences in the environment or nutrition upon the reproductive performance of the pigs. This means that limited extrapolation from these samples to the MS breed as a whole is advisable (Bidanel *et al.*, 1989).

There is evidence that the MS breed has larger litters through improvements in PS levels, and not through increased OR (Haley & Lee, 1993). Indeed, during the peri-implantation period (days 12-18 of gestation), the MS breed displays a significant reduction in the percentage of embryo lost compared with the losses seen in commercial European breeds such as the LW (Ford, 1997; Wilson *et al.*, 1999). Even when the uterus size and the OR of the two breeds has been observed to be similar, the MS breed has larger litters (Haley & Lee, 1993).

The MS displays an increased UC, achieved by a greater level of organisation in the uterus, as well as increased levels of placental efficiency compared to the US/European breeds. This increased UC helps them to increase LS through a reduction in embryo and foetal mortality. As UC becomes limiting, the number of viable conceptuses in a litter is no longer associated with OR and begins to be

associated with the average placental efficiency for that litter (Wu *et al.*, 1988; Vallet & Christenson, 1993; Wilson & Ford, 2001).

During the last third of gestation, together with its rapid growth, LW pigs show a significant increase in placental size and weight, but this does not occur in the highly prolific Chinese pig breeds, like MS (Zhou *et al.*, 2009). When foetal demands increase dramatically, the MS increases the density of placental blood vessels compared to the increase in placental size in LW (Ford, 1997; Mesa *et al.*, 2003). The vascular density of MS placenta increases between day 90 and 110 (Wilson *et al.*, 1998), which should increase the efficiency of nutrient and waste product exchange from maternal blood per unit area of placenta-endometrial interface to meet the demands during this period. However, the size of the placenta is kept constant for MS conceptuses, which have smaller placenta than conceptuses of US or European breeds (Ashworth *et al.*, 1990b; Wilson *et al.*, 1998; Biensen *et al.*, 1998). This decreased placental size allows MS females to accommodate more conceptuses in a similar uterine space to that in less prolific breeds (Bazer *et al.*, 1988). A possible explanation for the smaller placenta size in MS may be the reduced hystotroph secretion prior to elongation, which also slows conceptus development and results in smaller foetuses. The MS uterine environment restricts the development and elongation rate of the conceptus during the peri-implantation period and consequently allows a decreased prenatal losses and possibly increased UC (Anderson *et al.*, 1993; Ford & Youngs, 1993; Vallet *et al.*, 1998; Wilson *et al.*, 1998; Biensen *et al.*, 1999).

Studies performed in the uterine flushing of pregnant pigs have detected a reduction in endometrial secretion of proteins observed as a consequence of the reduced oestrogen biosynthetic capacity of the conceptus that causes the decreased growth rate in MS conceptuses (Ashworth *et al.*, 1990a; Anderson *et al.*, 1993; Ford & Youngs, 1993; Youngs *et al.*, 1994; Wilson *et al.*, 1998; Wilson *et al.*, 1999). Thus, this more gradual increase of secretion of oestradiol and proteins by the more-developed embryos in a MS litter may alter the uterine environment more gradually, increasing the probability of survival of less-developed littermates (Anderson *et al.*, 1993; Ford, 1997; Vallet *et al.*, 1998). In LW breeds, in contrast, the larger and more

advanced embryos alter the uterine environment to the detriment of the smaller embryos (Pope *et al.*, 1990). The oestrogen secreted as maternal recognition of pregnancy in pigs is oestradiol-17β (E2β). The expression of this oestrogen has been compared in MS and LW x Landrace (LR) conceptuses, finding differences in the time of expression between these two breeds (Pickard & Ashworth, 1995). In a later study by the same group (Pickard *et al.*, 2003), the expression of E2β in MS was found to be temporally determined and not conceptus stage dependent as in Western pig breeds. Thus, the size of the embryo does not determine the secretion of E2β in MS.

The differences in ES levels of MS and LW pigs has also been attributed to the lower variation in embryo length, weight, and distance between attachment sites (Wilmut *et al.*, 1992). At term, individual MS piglets are about 20-25% lighter than those from LW females and the lower variation in the distance between MS embryos avoids competition, reducing foetal mortality (Galvin *et al.*, 1993). Reduced growth rate has been found to be affected both by uterine environment and conceptus breed in embryo transfer studies in early stages of pregnancy (Ashworth *et al.*, 1990b; Ford & Youngs, 1993; Youngs *et al.*, 1994; Mesa *et al.*, 2003). This reduced size is also related with the E2β, which slow the growth of the embryo.

1.7. Selection for increase in litter size

Genetics plays an integral role in the control of the different reproductive traits such as OR, TN, GL, AP, UC, ES, and LS. Moreover, the genetic correlation for LS at different parities is very high, so successful selection on the first two parities should be effective for the subsequent parities (Avalos & Smith, 1987). The correlation between production and reproduction traits has been investigated in numerous studies with different results, from negative to positive or absence, always resulting in small genetic correlations (Rothschild, 1996; Sonesson *et al.*, 1998; Ruiz-Flores & Johnson, 2001; Noguera *et al.*, 2002; Foxcroft *et al.*, 2006; Rosendo *et al.*, 2010). Selection experiments to increase LS in mice have been successful (Bradford, 1968; Falconer, 1971). Thus, an increase in LS in swine through selection was expected to

be effective. However, direct selection for LS in swine has not been very successful (Bolet *et al.*, 1986; Avalos & Smith, 1987). However, previous studies have shown that OR, ES, and UC, components of LS, have high to moderate heritabilities (h^2 = 0.4, 0.3, 0.2, respectively), and respond well to long term selection (Bennett & Leymaster, 1989; Rothschild, 1996; Johnson *et al.*, 1999d). This has encouraged researchers to consider selection based on these components or their combination as an approach to increase LS.

Selection for increased OR in mice was directly effective but did not immediately result in increased LS (Bradford, 1968; Lamberson *et al.*, 1991). In the same way, OR responds to direct selection in swine, but the returns in terms of LS have been minimal (Cunningham *et al.*, 1979; Lamberson *et al.*, 1991). These results demonstrated that a single gene, hormonal treatment or nutritional regime would be unlikely to improve both OR and UC (Bennett & Leymaster, 1989). Greater response can be expected from selection for an index of OR and ES or UC than from direct selection in pigs for one of these traits (Bennett & Leymaster, 1989; Bennett & Leymaster, 1990; Cassady *et al.*, 2001).

Bennett & Leymaster (1989) proposed and extended a model for LS at birth (Bennett & Leymaster, 1990) that was dependent on OR, embryonic viability, and UC. Genetic analysis of the model suggested that selection for either of the components individually would not maximise the selection response for LS. Simulation of selection for LS and its components showed that LS increased most when selection was for indexes of OR and UC, followed by selection for indexes of OR and either LS or ES. Therefore, an index of OR and UC could lead to greater responses in LS than direct selection for LS (Bennett & Leymaster, 1990).

The University of Nebraska undertook 11 generations of selection for increases in an index of OR and ES rate, followed by three generations of selection (from generation 12 to 14) for LS or increased number of fully formed (NFF) pigs at parturition (Rathje *et al.*, 1997; Johnson *et al.*, 1999d; Mesa *et al.*, 2003). As a result, only approximately 50% of the increase in LS was in live piglets (Johnson *et al.*, 1999d). A range of studies have been performed in these selection lines to understand the

genetic and physiological control of reproduction. The detected increased LS at day 50 of gestation in the selected line was seen as contributing to greater foetal losses in late gestation, greater NSB, and lighter pigs at birth, leading to lower pre-weaning viability (Johnson *et al.*, 1999d). As a result, the improvement in LS in this selection experiment was similar to the one expected from direct selection for LS (Johnson *et al.*, 1999d). For future improvements in LS there is a need to gain a clearer understanding of the balance of follicular maturation and timing of ovulation, as well as embryonic, uterine, and placental factors associated with conceptus development and survival (Geisert & Schmitt, 2002). In view of the increased NMUM and NSB along with increased NFF pigs after selection for LS, traits related to foetal growth and development, such as weight of the live pigs, were detected as possible factors to be included in selection criteria (Johnson *et al.*, 1999d; Ruiz-Flores & Johnson, 2001)

The greater LS in MS pigs has been related to both smaller placenta and lighter foetuses (Christenson *et al.*, 1993; Biensen *et al.*, 1998). These characteristics suggested that selection for placental efficiency in pigs would decrease the placental size, and therefore more conceptuses could be accommodated in the same amount of uterine space (Wilson *et al.*, 1998; Biensen *et al.*, 1998). Wilson *et al.* (1999) found that selection for placental efficiency resulted in a correlated increase in LS. However, the sample employed was small and the selection was performed for just one generation. A more recent and larger selection experiment for four generations, using divergent selection for the foetal weight to placental weight ratio did not result in improvements in LS (Mesa *et al.*, 2005). The need to take into account other factors affecting weight and other measures of placental efficiency was suggested as a reason for the inefficiency of this selection method.

With the increase in LS produced by these direct or indirect selection methods, there is a consequently decrease in piglet growth and increase in piglet mortality (Johnson *et al.*, 1999d). For this reason, and in order to decrease piglet mortality, mothering ability is an important trait (Chen *et al.*, 2010). One of the traits related to the mothering ability is the TN and some selection has been performed in the pig

industry (Pumfrey *et al.*, 1980). Thus selection for these traits could improve the piglets outcome and the mothering ability indirectly (Hanenberg *et al.*, 2001).

As mentioned previously, the key porcine reproductive traits are expressed only in females and display low heritabilities (Avalos & Smith, 1987; Bennett & Leymaster, 1989) making the improvement of these traits by selective breeding difficult. These low heritabilities demonstrate the high influence of environmental factors in these traits, factors that can be partly controlled. However, the genetic effect on these traits, measured by the heritability, is low and thus, difficult to define but could benefit from the modern molecular techniques. Despite the relative failure of direct and indirect selection for LS, the pork industry has achieved tremendous gains in LS, through genetic selection and the introduction of hyperprolific dam lines into commercial production, along with improvements in nutrition, housing, and herd health management. However, by integrating molecular genetics approaches, such as MAS, into traditional selective breeding methods, it should be possible to achieve the maximum improvement in the economic value of domesticated livestock populations (Lande & Thompson, 1990). Thus, the finding of genes associated with these traits, together with the control of environmental factors and the high correlation between parities, will translate in a reproduction and production improvement in pigs.

The substantial difference in reproductive performance between pig breeds indicates that there is useful genetic variation available for genetic improvement. It has to be taken into consideration that quantitative traits, such as LS and OR, are usually affected by many genes, and consequently, the benefit from MAS is limited by the proportion of the genetic variance explained by a marker or gene (Meuwissen *et al.*, 2001).

Applications of MAS will improve as more associations between markers and traits are identified since LS is a multi-factorial trait (Cassady *et al.*, 2001). Besides, the use of MAS is expected to increase exponentially as genome sequencing projects increase the density of useful, segregating markers for economically important traits. The completion of the draft pig genome sequence and the development of high

density single nucleotide polymorphisms (SNPs) chips is allowing the performance of large scale association studies for pig reproductive traits (Bazer & Spencer, 2005).

1.8. Identification of genes controlling reproduction traits

Two approaches have been pursued to identify genetic markers for reproduction traits. First, genome scans employing anonymous DNA markers such as microsatellites have been used to identify quantitative trait loci (QTL) with effects on reproductive traits (Rathje *et al.*, 1997; Rohrer *et al.*, 1999; Wilkie *et al.*, 1999; Yasue *et al.*, 1999; Wada *et al.*, 2000; Rohrer, 2000; Cassady *et al.*, 2001; de Koning *et al.*, 2001; Hirooka *et al.*, 2001; King *et al.*, 2003; Lee *et al.*, 2003; Beeckmann *et al.*, 2003; Holl *et al.*, 2004; Rodriguez *et al.*, 2005; Sato *et al.*, 2006; Buske *et al.*, 2006a; Tribout *et al.*, 2008; Bidanel *et al.*, 2008; Li *et al.*, 2009; Ding *et al.*, 2009; Noguera *et al.*, 2009; Onteru *et al.*, 2011; Onteru *et al.*, 2012). Second, physiological candidate gene approaches, in which genes with known roles in the trait of interest are scanned for polymorphisms, which in turn are tested for associations with variation in the trait, have also been employed (Rothschild *et al.*, 1996; Short *et al.*, 1997; Rothschild *et al.*, 2000; Drogemuller *et al.*, 2001; Linville *et al.*, 2001; Jiang *et al.*, 2001; van Rens *et al.*, 2002; Gladney *et al.*, 2004; Buske *et al.*, 2005; Vallet *et al.*, 2005a; Campbell *et al.*, 2008; Zhou *et al.*, 2009; Lin *et al.*, 2009; Muñoz *et al.*, 2010; Fernandez-Rodriguez *et al.*, 2010; Fernandez-Rodriguez *et al.*, 2011).

The finding of a candidate gene with a role in reproduction mapped to the chromosomal location of a QTL enhances the possibility that this gene is really involved in the trait of interest. The QTL regions identified in genome scans are usually large, which makes it difficult to identify positional candidate genes. Furthermore, many genes are responsible for just a part of the phenotypic variation. This means that a combination of markers that capture most of the genetic variation for the trait will be needed for an improvement. Moreover, it is difficult to standardise all environmental influences for the long periods involved in studies of

reproductive performance. For these reasons, it is good to test the gene variants in several populations to detect general effects (Buske *et al.*, 2006a).

1.8.1. Linkage maps

The establishment of detailed linkage maps is a starting point for finding the chromosomal regions controlling phenotypic characteristics (i.e., QTL), and the subsequent identification of genes underlying the phenotypes of interest. The construction of linkage maps relies on recombination events between markers and following the segregation of marker alleles through the generations. Where a QTL for a particular trait is closely linked to these genetic markers, the alleles at that marker will appear to be associated with different levels of performance in the trait of interest.

Genetic markers are essential to the construction of linkage maps, and to map QTL. The markers should be abundant, informative, evenly distributed, highly polymorphic, and readily genotyped. Microsatellites were the markers of choice for QTL mapping studies, and they fulfil these criteria. A microsatellite consists of a sequence of 2 to 5 base pairs (bp) long repeated several times end to end, at specific sites throughout the genome and they are genotyped to determine the alleles of an individual. Nowadays, SNPs are used for large scale association studies covering the whole genome and allowing the discovery of more specific regions in the genome affecting the traits of interest, and enabling the discovery of candidate genes.

A number of low-resolution genetic maps of the pig genome have been published. One of the first linkage maps covering all 18 autosomes of the pig was published in 1995 (Archibald *et al.*, 1995), followed by a large map containing approximately 1,200 markers (Rohrer *et al.*, 1996). Numerous linkage analyses have been performed in order to map QTL (Ellegren *et al.*, 1994; Rohrer *et al.*, 1996; Marklund *et al.*, 1996; Rathje *et al.*, 1997; Wilkie *et al.*, 1999; Mikawa *et al.*, 1999; Cassady *et al.*, 2001; Bidanel *et al.*, 2001; Rodriguez *et al.*, 2005; Humphray *et al.*, 2007; Tribout *et al.*, 2008; Guo *et al.*, 2008; Noguera *et al.*, 2009; Vingborg *et al.*, 2009;

Guo *et al.*, 2009). These maps were primarily constructed on the basis of anonymous microsatellites and restriction fragment length polymorphism (RFLP) markers. One of the most recent maps for the whole genome contained not only microsatellite but also SNPs, some of which were used in genome-wide association studies (Vingborg *et al.*, 2009).

1.8.2. Quantitative trait loci analysis

The aim of a QTL analysis is the identification of genomic regions with effects on the trait of interest (phenotypic trait) (Falconer & Mackay, 1996; Kearsey, 1998). The statistical support for a QTL can be improved by typing additional markers over the same individuals, but the most efficient way to improve it, is to increase the number of animals in the study. Crosses between genetically and phenotypically diverse lines represent a powerful design for QTL mapping experiments (Rohrer *et al.*, 1994). If the lines differ widely in phenotype, it could be assumed that they are fixed for alternative alleles (Haley *et al.*, 1994; Rathje *et al.*, 1997). Under this assumption, all F_1 animals are expected to be heterozygous for many markers and many QTL.

The common procedure to search for QTL has been trait by trait. The traits, however, are often genetically correlated and result from a complex interaction of several different factors, and hence, the same QTL may affect two or more traits (Knott & Haley, 2000). Moreover, LS is affected by a set of factors such as farm, feed, season and mating boars, which account for a large part of the phenotype variance. From a genetic point of view, LS could be controlled by numerous genes in complicated physiological networks such as those affecting OR, fertilisation rate, ES and UC. Consequently, each QTL could explain just a small proportion of phenotype variance. Due to the importance of an improvement in production, the number of QTL studies is numerous, and not only for reproductive traits. Nowadays, the number of association studies using SNPs has increased due to the availability of a SNP chip (Illumina PorcineSNP60 BeadChip) which allows the genotyping of 60,000+ SNPs simultaneously for each pig (Ramos *et al.*, 2009) . These studies will

help dissect the genetic control of a range of important traits including reproductive performance.

1.8.2.1. Previous QTL studies for reproductive tratis

The number of QTL with effects on reproductive traits which have been identified so far is more limited than those with effects on other production traits. Only few QTL affecting LS have so far been reported in the literature. Moreover, most of these QTL were identified in crosses between selected lines (Cassady *et al.*, 2001) or in crosses involving the prolific MS breed (Table 1.1). However, further research is required to find the causative genetic variation in the gene influencing the trait variation within breeds to use in MAS programmes.

Most of the QTL found in published studies are annotated in the pigQTLdb (www.animalgenome.org/QTLdb/pig.html), where for each QTL there is information for the trait, the chromosome, the position, the population, and publication. The reasons for variations in results between studies could be many, such as differences between resource populations, number of evaluated animals, mating systems, definition and measurement of the phenotypic traits and environmental influences (Buske *et al.*, 2006a). ES is calculated by dividing the NVE at day 30 of gestation by the OR on dissected ovaries, thus, the gilt or sow is slaughtered at this stage in order to measure the traits. In contrast, PS is calculated by dividing TBA by the OR estimated by laparoscopy. The NSB are the piglets that die just before, during or soon after farrowing. The NMUM indicate the number of piglets that die in uterus and for which degradation starts in the uterus. These piglets reflect the losses after bone formation starts, since embryos that die prior to this development stage will be reabsorbed in the uterus.

Trait	SSC	Population	Reference
Age at puberty	1, 10	WC x MS	(Rohrer *et al.*, 1999)
	7, 8, 12	LW x LR	(Cassady *et al.*, 2001)
	7, 8, 12, 15	LW x LR	(Holl *et al.*, 2004)
	1, 4, 6, 13	MS x LW	(Bidanel *et al.*, 2008)
	1, 7, 8, 17	D x Er	(Yang *et al.*, 2008)
Ovulation rate or	4, 8, 13, 15	LW x LR	(Rathje *et al.*, 1997)
Number of corpora lutea	7, 8, 15	Y x MS	(Wilkie *et al.*, 1999)
	3, 4, 8, 9, 10, 13, 15, X	WC x MS	(Rohrer *et al.*, 1999)
	8	Y x MS	(Braunschweig *et al.*, 2001)
	9	LW x LR	(Cassady *et al.*, 2001)
	8	MS x LW	(Jiang *et al.*, 2001)
	8	MS x WC	(Campbell *et al.*, 2003)
	9	LW x LR	(Holl *et al.*, 2004)
	3	MS x D	(Sato *et al.*, 2006)
	4, 5, 7, 9, 13	MS x LW	(Bidanel *et al.*, 2008)
Embryo survival	9, 12, 18	MS x LW	(Bidanel *et al.*, 2008)
Prenatal survival	8	MS x LW	(King *et al.*, 2003)
Uterine capacity	8, X	WC x MS	(Rohrer *et al.*, 1999)
Uterine length	5, 7	Y x MS	(Wilkie *et al.*, 1999)
Gestation length	1, 9, 15	Y x MS	(Wilkie *et al.*, 1999)
	8	MS x LW	(Jiang *et al.*, 2002b)
	6, 11	Ib x MS	(Casellas *et al.*, 2008)
	2, 8, 12	D x Er	(Chen *et al.*, 2010)
Litter size or Total number born	1	WC x MS	(Rothschild *et al.*, 1996)
	6	GMP x MS	(Yasue *et al.*, 1999)
	6	Y x MS	(Wilkie *et al.*, 1999)
	7, 12, 14, 17	LW/LR x MS	(de Koning *et al.*, 2001)
	11	LW x LR	(Cassady *et al.*, 2001)
	8	LW x MS	(King *et al.*, 2003)
	11	LW x LR	(Holl *et al.*, 2004)
	1	LW	(Horogh *et al.*, 2005)

	7	(LW x LR) x Lc	(Buske *et al.*, 2005)
	7, 15	WD x Er	(Li *et al.*, 2009)
	13, 17	Ib x MS	(Noguera *et al.*, 2009)
	1	WC x MS	(Rothschild *et al.*, 1996)
	12	Ib x MS	(Fernandez-Rodriguez *et al.*, 2010)
	7	(LW x LR) x Lc	(Buske *et al.*, 2005)
Number born alive	1	WC x MS	(Rothschild *et al.*, 1996)
	11	LW x LR	(Cassady *et al.*, 2001)
	11	LW x LR	(Holl *et al.*, 2004)
	7	(LW x LR) x Lc	(Buske *et al.*, 2005)
	1	LW	(Horogh *et al.*, 2005)
	12, 13, 17	Ib x MS	(Noguera *et al.*, 2006)
	1	(LW x LR) x Lc	(Buske *et al.*, 2006b)
	7, 16, 18	LW x F LR	(Tribout *et al.*, 2008)
	6, 15	WD x Er	(Li *et al.*, 2009)
	13, 17	Ib x MS	(Noguera *et al.*, 2009)
	12	Ib x MS	(Fernandez-Rodriguez *et al.*, 2010)
	2, 14	(MS x J) x LW	(Muñoz *et al.*, 2010)
Number of still born	4, 5	Y x MS	(Wilkie *et al.*, 1999)
	5, 13	LW x LR	(Cassady *et al.*, 2001)
	7, 12, 14, 17	LR x LW, LW	(Andersson & Georges, 2004)
	5, 12, 13, 14	LW x LR	(Holl *et al.*, 2004)
	6, 11, 14	LW x F LR	(Tribout *et al.*, 2008)
	7, 8	WD x Er	(Li *et al.*, 2009)
Number of viable embryos	6, 9, 12, 18	MS x LW	(Bidanel *et al.*, 2008)
Total mummified	2, 6, 12	LW x LR	(Holl *et al.*, 2004)
Teat number	1, 3, 10	WC x MS	(Rohrer, 2000)
	1, 3, 7	GMP x MS	(Wada *et al.*, 2000)
	1, 6, 7, 8, 11	LW x LR	(Cassady *et al.*, 2001)

2, 3, 10, 12	MS x LW/LR	(Hirooka *et al.*, 2001)
8	LW x MS	(King *et al.*, 2003)
1, 8	MS, P, WB crosses	(Beeckmann *et al.*, 2003)
2, 5	MS, P, WB crosses	(Lee *et al.*, 2003)
1, 6, 7, 8, 11, 15	LW x LR	(Holl *et al.*, 2004)
5, 10, 12	Ib x MS	(Rodriguez *et al.*, 2005)
3, 8, 12	MSx D	(Sato *et al.*, 2006)
6, 7	MS x Y	(Zhang *et al.*, 2007)
3, 4, 7, 8, 11, 16	LW x MS	(Bidanel *et al.*, 2008)
1, 3, 4, 6, 7, 10, 11, 12, 16, 17	MS x LW	(Guo *et al.*, 2008)
1, 3, 4, 5, 6, 7, 8, 12	WD x Er	(Ding *et al.*, 2009)

Table 1.1 QTL for female reproductive traits in pigs. The Table indicates the trait, the chromosome number where the QTL was mapped (SSC), the population used in the study and the reference of the study. D, Duroc; Er, Erhualian; F LR, French Landrace; GMP, Gottingen miniature pig; Ib, Iberian; J, Jiaxing; LW, Large White; LR, Landrace; Lc, Leicoma; MS, Meishan; P, Pietran; WC, white composite; WB, Wild boar; WD, White Duroc; Y, Yorkshire.

Recently, a whole-genome association study using the Porcine SNP60 Beadchip (Ramos *et al.*, 2009) was conducted for pig life-time reproductive traits, in a LW and a LW x LR population (Onteru *et al.*, 2011). Another study by the same group (Onteru *et al.*, 2012), conducted a whole-genome association study for reproductive traits, which included LS, TBA, NSB, NMUM and GL, in the same population as the previous study. A large number of regions were associated with the different traits in the different parities, and the results were compared with previous QTL studies. Different QTL regions were detected for the three different parities for each trait. Some of the chromosomal regions identified in this study had not previously been identified as QTL or as QTL with effects on reproductive traits. The results from this first use of the power of genome-wide studies using 60,000+ SNPs indicate not only the potential of such studies to find associations but also the challenge of exploring multiple regions for candidate genes and causal genetic variants.

The focus on chromosome 8 in the current study builds upon findings at The Roslin Institute (Jiang *et al.*, 2002b; King *et al.*, 2003). On porcine chromosome 8 (Sus scrofa chromosome 8-SSC8), several QTL for female reproductive traits have been identified, a QTL for UC at the 71 cM position (Rohrer *et al.*, 1999) and QTL for LS and PS at position 127 cM and 125 cM, respectively (King *et al.*, 2003). For OR different QTL were found on SSC8, at 107.5 cM (Rathje *et al.*, 1997), 5cM (Rohrer *et al.*, 1999; Campbell *et al.*, 2003), 99 cM (Braunschweig *et al.*, 2001), 80 cM (Jiang *et al.*, 2002b) and at 101 cM (Wilkie *et al.*, 1999). Recently, Li *et al.* (2009) found a suggestive QTL at 84 cM for NSB. Also QTL for TN were found at 16.2 cM (Cassady *et al.*, 2001), 20 cM (Holl *et al.*, 2004) 47 cM (King *et al.*, 2003), 29-46 cM, 56-74 cM (Sato *et al.*, 2006), 63.3 cM (Beeckmann *et al.*, 2003) and 94 cM (Bidanel *et al.*, 2008), and recently by Ding *et al.* (2009) at 86.4 cM. As many different markers and linkage maps have been used in these studies the QTL positions cited should be treated as indicative rather than directly comparable. QTL for AP (Cassady *et al.*, 2001; Holl *et al.*, 2004; Yang *et al.*, 2008) and GL (Jiang *et al.*, 2001; Chen *et al.*, 2010) were also mapped on this chromosome.

1.8.3. Candidate genes

After identification of a QTL, regions that usually cover 10-20 cM, the ultimate goal is to identify the responsible gene itself, and the causative mutation for the phenotypic variation. The first steps toward this challenging aim is the fine mapping of the QTL and merging of the mapped QTL with candidate genes in this chromosomal region. A gene can be suggested as a potential candidate gene for reproduction because of the important physiological role it plays in reproduction (**physiological candidate genes**). Alternatively or additionally, candidate genes can be chosen on the basis that they are genes that map to the putative QTL region (**positional candidate genes**). By using information on orthologous genes in homologous chromosomal regions of other species (**comparative positional candidate genes**) can be identified. Finally, genes which are differentially expressed between individuals in the tissue under investigation can also be consider as candidate genes. Therefore, one requirement for the finding and investigation of

candidate gene expression is the material to test, i.e., a representative sample and a sufficient number of individuals is necessary in order to detect the effect of gene variants in different populations.

A marker associated with a trait in one population of pigs may not be associated with that trait in other populations. However, polymorphisms that alter gene functions would be expected to have similar effects across different populations of pigs (Vallet *et al.*, 2005a). Mutations in the coding regions of a gene may change the function or functionality of the encoded protein. Mutations in regulatory regions of a gene that affect its expression (level, timing or tissue-specificity) can change the availability of protein of normal function (Buske *et al.*, 2006a). Thus, when a marker is linked to a genetic variation in a trait, the candidate gene could contain the causal mutation mapped by the marker or it could be just linked to the marker in a regulatory way, with the marker position close or not to the candidate gene and its genetic variation.

The proportion of pig genes that have been mapped is small; consequently, the number of positional candidate genes is limited. However, a draft pig genome sequences have been released and are accessible in genome browsers such as Ensembl (http://www.ensembl.org/Sus_scrofa/Info/Index). The Swine Genome Sequencing Consortium has recently deposited the sequence assembly (Sscrofa10.2) on which the pig genome sequence publications will be based (Archibald *et al.*, 2010). The genome annotation currently accessible through the Ensembl genome browser is based on an earlier less complete genome sequence assembly (Sscrofa9). Therefore, as information on the gene content of the pig genome is incomplete it is useful to consider predictions based on comparative genome mapping.

Despite the extensive conservation of genome sequence and chromosomal organisation that exists between mammalian genomes, gene order and distance differ between species. Therefore, comparative mapping is critical to identify those chromosomal segments conserved during evolution and their rearrangements in the different species (Nadeau & Sankoff, 1998). Comparative genetic maps indicate that there is more structural similarity between pigs and human than, for example, mouse

and human (Humphray *et al.*, 2007). For example, Human chromosome 4 (Human Homo sapien chromosome 4-HSA4) shares extensive homology with SSC8 as well as with SSC15 and SSC17, but gene orders differ between HSA4 and SSC8.

The oestrogen receptor (*ESR*) gene, which is located on SSC1, encodes a steroid binding hormone receptor and is associated with increased LS. The association of *ESR1* with LS was first reported by Rothschild *et al.* (1996) who found a PvuII polymorphism in intron 9 of *ESR1* in different populations: MS, MS synthetic lines and LW populations. The *ESR1* B allele, mediating the actions of oestrogens, has been associated with increased LS and TBA in several studies (Rothschild *et al.*, 1996; Short *et al.*, 1997; Isler *et al.*, 2002; Horogh *et al.*, 2005). On the contrary, some much smaller studies reported an association of superior LS with the A allele rather than the B allele (van Rens *et al.*, 2002; Goliasova & Wolf, 2004). In addition, no significant association of *ESR1* with LS was detectable in some swine populations (Drogemuller *et al.*, 2001; Gibson *et al.*, 2002; Muñoz *et al.*, 2010).

The retinol-binding protein 4 (*RBP4*) gene on SSC14 has been suggested as a candidate gene for LS based on its role. RBP4 protein provides the conceptus with appropriate amounts of retinoic acid in the early critical phase of pregnancy around day 12 (Rothschild *et al.*, 2000). Retinoic acid is implicated in the regulation of gene transcription and trophoblast elongation (Harney *et al.*, 1990). Allele effects that differ between lines have been found in some studies (Rothschild *et al.*, 2000; Linville *et al.*, 2001) but not in all (Drogemuller *et al.*, 2001).

Bone morphogenetic protein receptor 1B (*BMPR1B*) is a member of the transforming growth factor-β (*TGF-β*) receptor family that plays a pivotal role in bone formation during embryogenesis and fracture repair. The *BMPR1B* system plays an important physiological role in the regulation of ovarian function as well as oocyte development (Shimasaki *et al.*, 1999). In addition, a non-conservative substitution in *BMPR1B* coding sequence is found to be fully associated with the hyper prolificacy phenotype of Booroola ewes. Because of its known effect on OR in sheep, *BMPR1B* was considered a candidate gene for LS on SSC8. In swine, *BMPR1B* does not seem

to be a major gene for prolificacy associated with LS, OR or PS (Kim *et al.*, 2003; Tomas *et al.*, 2006; Casellas *et al.*, 2008).

The following genes have also been suggested as candidates for reproductive traits. Prolactin receptor (*PRLR*) gene plays a role in the maintenance of gravidity but no significant associations of the *PRLR* locus with litter and growth performance traits were detected (Drogemuller *et al.*, 2001; Linville *et al.*, 2001). SPARC-like protein 1 or high endothelial venule protein (*SPARCL1*) has a role in cell adhesion and it is expressed in the ovary and placenta. The association of the erythropoietin receptor (*EPOR*), which controls the terminal differentiation and number of foetal red blood cells, with LS was investigated in a population of Yorkshire, Landrace and Duroc crossbred pigs, but no significant effect was found (Vallet *et al.*, 2005a). The epidermal growth factor (*EGF*) was investigated based on its role in the physiology of reproduction. EGF is produced by the conceptus and in the uterus of the sow (Mendez *et al.*, 1999). In the foetus, EGF stimulates growth and proliferation of skin epithelia to grow and mature (Gladney *et al.*, 1999). Properdin (*BF)* gene has an integral role in influencing uterine epithelium growth. *BF* was investigated in a commercial pig cross population (Buske *et al.*, 2006a). Expression of the gonadotrophin releasing hormone receptor gene (*GNRHR*) is critical in the endocrine regulation of reproduction, important for ovulation (Rohrer *et al.*, 1999; Jiang *et al.*, 2001). However, none of the QTL mapped for OR on SSC8 lie on the region of the chromosome where *GNRHR* is located, thus undermining the case for *GNRHR* as a candidate gene.

Trait	Associated gene	Gene name	SSC	Population	Reference
Age at puberty	PAX5	Paired box 5	1	D x BT and LR x BT	(Kuehn *et al.*, 2007)
	AKR1C2	Aldo keto reductase 1C2	10	¼ MS	(Nonneman & Rohrer, 2003)
	PRLR	Prolactin receptor	16	LW x MS	(van Rens & van der Lende, 2002)
Ovulation rate	NCOA1	Nuclear receptor coactivator 1	3	MS x LW	(Melville *et al.*, 2002)
	GNRHR	gonadotropin releasing hormone receptor	8	MS x LW	(Jiang *et al.*, 2001)
	MAN2B2	Mannosidase 2B2	8	MS x WC	(Campbell *et al.*, 2008)
	PRLR	Prolactin receptor	16	LW x MS	(van Rens *et al.*, 2003)
Uterine length	FSHB	Follicle stimulating hormone beta	2	LW x MS	(Lin *et al.*, 2009)
	PRLR	Prolactin receptor	16	LW x MS	(van Rens *et al.*, 2003)
Uterine capacity	EPOR	Erythropoietin receptor	2	Y x LR x CW x LW	(Vallet *et al.*, 2005b)
	STE	Oestrogen sulfotransferase	8	WC x MS	(Kim *et al.*, 2002)
	sFBP	Secreted folate binding protein	-	MS x W	(Vallet *et al.*, 2005a)
Teat number	ESR1	Oestrogen receptor 1	1	LW	(Short *et al.*, 1997)
	APOB	Apolipoprotein B-100	3	MS x D	(Sato *et al.*, 2006)
Litter size, TNB, NBA	ESR1	Oestrogen receptor 1	1	MS x SL and LW	(Rothschild *et al.*, 1996)
				LW	(Short *et al.*, 1997)
				CB	(Chen *et al.*, 2000)
				MS x LW	(van Rens *et al.*, 2002)
				LW	(Horogh *et al.*, 2005)
	ESR2	Oestrogen receptor 2	1		(Buske *et al.*, 2006b)
				Ib	(Muñoz *et al.*, 2004)
	FSHB	Follicle stimulating hormone beta	2	Y x EL	(Li *et al.*, 1998)
				LW x MS	(Li *et al.*, 2008)
	FUT1	fucosyl transferase 1	6	PBP	(Horak *et al.*, 2005)
	RNT4	ring finger protein 4 gene	6	CQ	(Niu *et al.*, 2009)

LEPR	Leptin receptor gene	6	Y, D	(Chen *et al.*, 2004b)
BF	Properdin	7	(LW x LR) x Lc	(Buske *et al.*, 2005)
SPP1	Secreted phosphoprotein 1	8	SL	(Korwin-Kossakowska *et al.*, 2002) (King *et al.*, 2003)
LIF	Leukemia inhibiroty factor	8	GL	(Spotter *et al.*, 2009)
			LW	(Lin *et al.*, 2009)
RBP4	Retinol binding protein 4	14	SL	(Rothschild *et al.*, 2000)
			GW	(Spotter *et al.*, 2009)
			MS, J x F LW	(Muñoz *et al.*, 2010)
IGF2	Insulin-like growth factor 2	14	MS, J x F LW	(Muñoz *et al.*, 2010)
PRLR	Prolactin receptor	16	LW	(Vincent *et al.*, 1998)
			MS, LR	(Vincent *et al.*, 1998)
			SL	(Drogemuller *et al.*, 2001)
			LW x MS	(van Rens *et al.*, 2003)
LEP	Leptin gene	18	SL	(Drogemuller *et al.*, 2001)
			Y; LR; D	(Chen *et al.*, 2004a) 2

Table 1.2 Candidate genes associated with female reproductive traits in pigs. The Table indicates the trait, the associated gene symbol and the name, the chromosome where it have been mapped (SSC), the population used in the study of the gene as a candidate gene and the reference. BT, Yorkshire x maternal Landrace composite; CB: Chinese breeds; CQ, Chinese Qingping; CW, Chester White; D, Duroc; Er, Erhulian; F LW, French Large White; G LR, German Landrace; G LW, German Large White; Ib, Iberian; J, Jiaxing; Lc, Leicoma; LW, Large White; PBP, Přeštice Black-Pied; LR, Landrace; MS, Meishan; SL, Synthetic lines; WC, white composite; W: White European breed cross; Y: Yorkshire.

In a recent study, Fernandez-Rodriguez *et al.* (2011) compared gene expression from ovaries from sows of different prolificacy levels during pregnancy ((IbxMS) x LW) and identified 27 candidate genes for all published QTL through the combination of microarray results and linkage analysis.

1.8.4. Secreted Phosphoprotein 1 (SPP1)

Secreted phosphoprotein 1 (*SPP1*, osteopontin, bone sialoprotein 1 and early T-lymphocyte activation factor 1) is a highly phosphorylated acidic glycoprotein member of the small integrin-binding ligand, N-linked glycoprotein (SIBLING) family of genetically related ECM proteins. These proteins are recognised as key players in diverse processes, such as bone mineralisation, cancer metastasis, cell-mediated immune responses, inflammation, and angiogenesis (Johnson *et al.*, 2003c). Thus, *SPP1* has potential to influence tissue remodelling at the conceptus-maternal interface by affecting cell-cell and cell-ECM communication, increasing cell proliferation, migration, and survival, and regulating local cytokine networks (Senger & Perruzzi, 1996; Johnson *et al.*, 2003c).

SPP1 undergoes extensive posttranslational modifications believed to be important to its function. These modifications include proteolytic cleavage, phosphorylation, glycosylation (Zhang *et al.*, 1992). Originally isolated from bone, SPP1 has been found in epithelial cells and in secretions of the gastrointestinal tract, kidneys, thyroid, breast, uterus, placenta, testis, leukocytes, smooth muscle cells, and highly metastatic cancer cells (Senger & Perruzzi, 1996; Johnson *et al.*, 1999b). Its expression is upregulated during the initial stages of pregnancy in uterus of pigs (White *et al.*, 2005) and other mammalian species, including humans (Johnson *et al.*, 2003a), mice, rabbits, goats, rats and sheep (Garlow *et al.*, 2002; Allan *et al.*, 2007), regardless of placental structure. A global gene profiling experiment using high-density microarray technology indicates that SPP1 is the most highly up-regulated ECM-adhesion molecule in the human uterus, as it becomes receptive to implantation (Kao *et al.*, 2002; Carson *et al.*, 2002).

The temporal *SPP1* mRNA expression is coordinated with the morphological and biological changes that conceptuses undergo during pregnancy recognition and early adhesion (implantation and placentation) between trophoblast and uterine LE (Garlow *et al.*, 2002). Both oestrogen and progesterone influence uterine-conceptus interactions during day 14 to 26 of pregnancy (Garlow *et al.*, 2002). *SPP1* is induced

in response to conceptus oestrogen. Progesterone, the hormone of pregnancy, supports *SPP1* expression in the endometrial GE that is associated with increasing production of histotroph required for foetal/placental development and growth (Garlow *et al.*, 2002; White *et al.*, 2005).

In contrast with humans and sheep, pigs express the *SPP1* gene directly in the luminal epithelium (LE) (Leiser & Dantzer, 1988) beginning on day 12, just before conceptus attachment (peri-implantation period), whereas *SPP1* mRNA is not induced in the glandular epithelium (GE) until between days 30 and 35 of pregnancy (Garlow *et al.*, 2002). Expression is then maintained in both LE and GE throughout gestation with a 20-fold increase in the GE.

In pig, an increased expression of *SPP1* mRNA has been shown to result in integrin activation, and the accumulation of the cytoskeletal molecules required to form the "focal adhesions" for adhesion and signalling between the conceptus and the uterus (Garlow *et al.*, 2002). SPP1 contains an Arg-Gly-Asp (RGD) sequence that mediates binding to cell surface integrin receptors, including αvβ3, α5β1, αvβ1, αvβ5, αvβ6 and α8β1 (Johnson *et al.*, 2003b). Alternative binding-sequence interactions between SPP1 and integrins such as α4β1, α9β1, α4β7 can also occur (Johnson *et al.*, 2003b). The expression of αvβ3 integrin at the apical surface of LE and αvβ6 on trophectoderm have been described to mediate attachment for implantation in pigs (White *et al.*, 2005; Erikson *et al.*, 2009). Moreover, β3 integrin is expressed during the putative implantation window in endometrial epithelial cells in human embryos (Campbell *et al.*, 1995) and in mouse embryos during the peri-implantation phase (Sutherland *et al.*, 1993).

The *SPP1* gene, with a key role in conceptus implantation and maintenance of pregnancy (Hao *et al.*, 2008), is a strong candidate gene as it is located under the peak of the SSC8 QTL with effects on PS (King *et al.*, 2003). *SPP1* is of interest as a mediator of successful pregnancy (White *et al.*, 2005). *SPP1* in the pig, which experiences significant conceptus loss, offers an excellent model to study these mechanisms because of its regulated and temporal pattern of LE and GE expression

during implantation and development critical periods (Garlow *et al.*, 2002; White *et al.*, 2005; Erikson *et al.*, 2009; Bailey *et al.*, 2010).

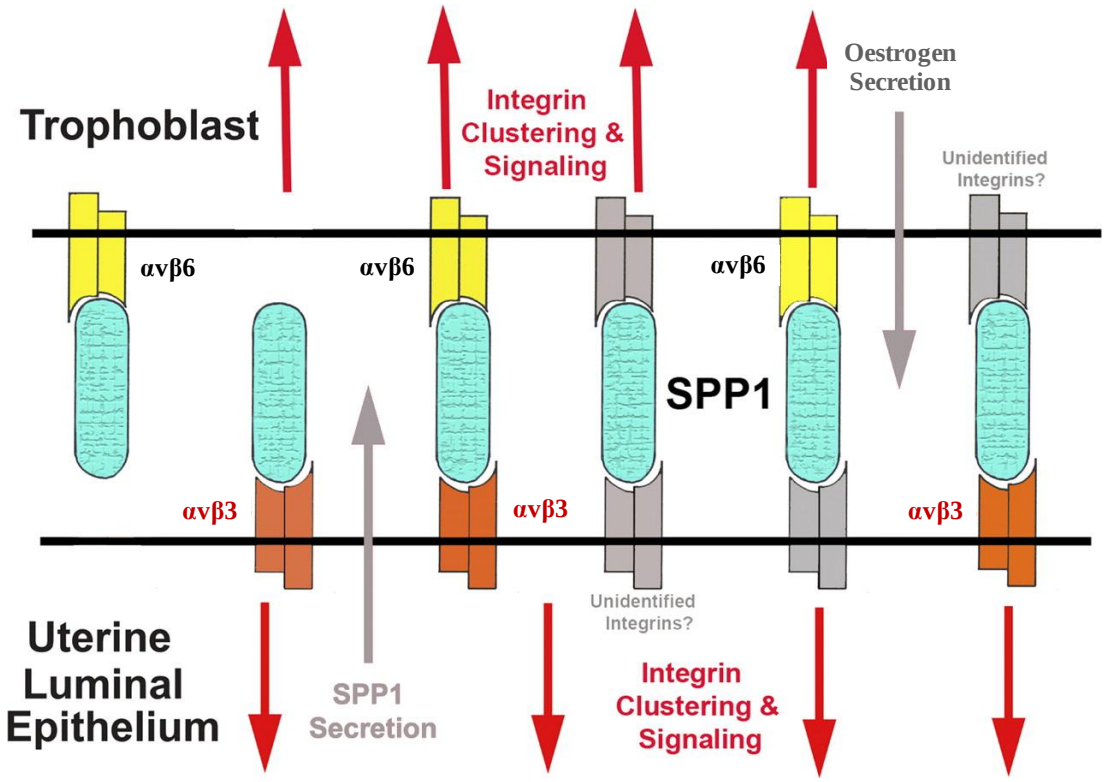

Figure 1.2 Summary of potential SPP1-intergrin interactions at conceptus-maternal interface in pigs at implantation. SPP1 binds to LE cells via αvβ3 and to the conceptus Trophoblast cells via αvβ6. There is the possibility that other as-yet-uninvestigated integrins, or other receptors, bind SPP1 on conceptus Tr and LE during the peri-implantation period of pigs. Adapted from Erikson *et al.* (2009).

1.9. Functional genomics

Bridging the gap between identifying a gene sequence and then determining its physiological role within an organism represents one of the greatest challenges of modern biology (Blomberg & Zuelke, 2004). Thus, greater understanding of the control of reproductive traits will require broad evaluation at the DNA, mRNA, protein, and detailed phenotypic levels, using a wide variety of techniques including DNA sequencing, evaluation of gene expression, and even mutational and transgenic

analysis, all of which are united under the term 'functional genomics' (Pomp *et al.*, 2001). An important step in examining functions of genes is to determine their spatial and temporal expression patterns in different tissues or under different conditions. The evaluation of the gene at these levels is the ultimate step in a QTL analysis.

For many years, researchers have focused on analysis of expression of individual genes involved in pig reproductive processes. The progress made with these studies has been slow. Therefore, modern techniques accessing many genes in parallel, such as microarray-based expression profiling, could benefit these studies. This approach not only applies to the expression studies but also to the gene studies where genome-wide and system-wide experimentation is required due to the complex nature of genetic control over polygenic traits, such as OR and LS (Pomp *et al.*, 2001). The pig transcriptome has been analysed to address biomedical, agricultural, and fundamental biological questions, using more and more sensitive and comprehensive tools (Tuggle *et al.*, 2007). Finally, comprehensive genome sequence annotation will allow rapid integration of gene expression data with gene sequences, linkage mapping and genome-wide association studies (Tuggle *et al.*, 2007).

Therefore, as mentioned previously, the selection for an increase in LS could benefit from the use of marker-assisted selection for what genes with important effects on quantitative traits need to be identified (Rothschild, 1996; Spelman & Bovenhuis, 1998; Davis & DeNise, 1998; de Vries *et al.*, 1998; Rothschild, 2004). One of the gene identified as a candidate gene for LS in pigs, ESR, has been used in marker-assisted selection since 1994 at PIC in the United States and Europe (Short *et al.*, 1997). The potential value of fixing the beneficial allele of ESR was estimated to be over $20 per sow per year (Short *et al.*, 1997).

1.10. Objectives

The aims of this study were to identify QTL affecting reproduction traits in pigs, and characterise candidate gene(s) underlying the genetic trait. This study was divided into two approaches.

1.10.1. Genetics approach: identification of QTL with effects on litter size and embryo survival

The initial objective concerning this approach was to map LS and PS QTL in pigs. Therefore, The Roslin LW x MS QTL mapping population, with records for female reproductive performance data including OR and LS, was genotyped for multiple genetic markers across the genome, with a particular focus on SSC8 where a previous study mapped QTL and a candidate gene (*SPP1*). QTL analyses are in essence tests for associations between variation in the trait and the genotypes across the genome, allowing the recognition of the inheritance of the chromosomal segments from the parental animals.

The second part of the genetic analysis involved the identification of positional candidate gene. For this objective, the new linkage maps for the pig chromosomes, and the improved map for SSC8 were aligned with the emerging draft pig genome sequence and with the annotated homologous regions in the human genome, and inspected for potential positional and physiological candidate genes. This exploited the genome conservation between mammalian genomes, in which gene content was expected to be very similar but the gene order can be different.

1.10.2. Physiology approach: functional characterisation of candidate genes.

The characterisation of positional and physiological candidate genes included functional characterisation such as determination of *SPP1* mRNA and protein

abundance in the endometrium and placenta a) between feto-placental units of different size occupying the same uterus and b) in different genotypes with clear differences in LS. This functional characterisation in relevant tissue samples involved RT-qPCR, *in situ* hybridization (ISH) and immunohistochemistry (IHC) studies for the quantification and location of mRNA and protein.

1.10.3. Hypothesis for *SPP1* as a candidate gene for reproductive traits

Having in mind the low success of increase in LS due to the associated losses, the approach in this study was focussed on the efficiency of MS breed during pregnancy. Therefore, the weight of the piglets at birth, determine during early pregnancy, was considered an important factor determinant for LS. Thus, it was hypothesised that the differences in the foetal growth, determinant for LS and PS, may be associated with conceptus attachment.

Chapter 2
Genome-wide linkage analysis

2.1. Introduction

2.1.1. Reproductive traits in pigs

Porcine female reproductive performance traits have low heritability and the fact that they are expressed only in females, limits improvement of these traits through traditional selective breeding programmes. However, there is abundant evidence for genetic variation that could be exploited to improve performance. Besides, this low heritability demonstrates a high impact of environmental factors in these traits, which improvement is limited but possible (Prunier *et al.*, 2010). Indeed, some success in selecting for increased LS and more recently for LS at day 5 has been achieved (Su *et al.*, 2007). An understanding of the genetic control of reproductive performance, a critical component of sustainable animal production systems, would offer the opportunity to utilise natural variation and improve selective breeding programmes through MAS.

The Chinese MS, a member of the Taihu group of breeds, is one of the most prolific pig breeds known; farrowing between three to five more live piglets per litter than the European commercial breeds, such as LW. However, the MS is not commercially viable in Europe due to its poor growth rate and high carcass fat content (Bidanel *et al.*, 1990; Haley *et al.*, 1992; Serra *et al.*, 1992). An understanding of the causes and regulation of this prolificacy would not only be of value from a scientific point of view, but might also indicate ways in which the prolificacy of other pig breeds could be enhanced. However, only a fraction of the molecular basis of this superior reproductive performance of the MS has been identified to date and it is evident that it merits further investigation.

The MS breed has larger litters through improvement in PS at a given OR and not through increased OR (Haley & Lee, 1993). With respect to OR, there is a discrepancy between studies in the UK and the US. In the UK studies, OR was found to be similar in gilts at comparable numbers of post-pubertal oestrous cycles, but in older sows, MS had a higher OR than LW pigs. However, the US studies showed

higher OR in multiparious MS sows than in multiparous composite white sows (1/4 each of Chester White, Landrace, Large White, and Yorkshire), but MS gilts do not always have a higher OR than composite white gilts. Generally, when the same number of cycles after puberty is compared, the OR is similar in MS and composite white and LW gilts. But when the comparison is made at the same age, then OR is higher in MS, possibly as a result of the earlier onset of puberty in this breed (Bazer *et al.*, 1988). These differences between breeds seem to increase as the sows get older (Christenson *et al.*, 1987; Bennett & Leymaster, 1989; Haley & Lee, 1993). Uterine size has been observed to be similar between breeds (Haley & Lee, 1993). However, the MS breed has been shown to display an increased UC, achieving this by a greater level of organisation in the uterus (Christenson *et al.*, 1987; Haley & Lee, 1993), as well as increased levels of placental efficiency compared to both European and US breeds.

2.1.2. Genetics of reproduction

Genetics plays an integral role in the control of reproductive traits such as OR, TN, GL, AP, testicular size, UC, ES, PS and LS parameters, comprising the TNB, the TBA and NSB. Measuring LS as a reproductive trait is straightforward and it is the most important trait for pig producers. More recently, in response to increases in perinatal mortality, Danish pig breeders have been selecting for the number of pigs alive five days after birth and have made progress with this breeding objective (Su *et al.*, 2007). Currently, LS varies from approximately 2 to 20 pigs per litter, with means from 9 to 11, depending on the breed. By increasing the NVE per litter that an individual sow farrows, the size of the population of breeding females can be reduced yielding a more efficient production system with increased outputs and reduced overhead costs and environmental footprints.

Since heritability estimates for these reproductive traits are usually 0.10 or less, the response to direct selection would be expected to be low. However, LS is predicted to increase by improving any of its components such as OR and PS, when the other components are not limiting (Bennett & Leymaster, 1989; Mesa *et al.*, 2003), and

heritability estimates for these traits appear to be higher than for LS. Selection for traits, such as OR and ES, however, is not practical in a commercial environment, due to the aggressiveness of the measure technique, laparoscopy, which due to the anatomy of the pig is the only technique efficient to measure OR. There have been a number of selection studies, not only directly for LS but also for related traits, such as OR, UC or ES. Johnson *et al.* (1999d) found increases of approximately 7.4 ova and 3.8 foetuses at day 50 of gestation and 2.3 fully formed pigs and 1.1 live pigs at birth after 11 generations of selection using an index for increased OR and ES. The response after a further three generations of selection for LS was an additional three fully formed pigs and 1.4 live pigs per litter. In another study, Gama & Johnson (1993) tested the results of eight generations of selection for LS and observed an increase of a 1.21 ± 0.38 pigs. As a result of these selection studies and the use of superior breeds and lines as well as crossbreeding, LS per sow in the US has increased from 6.0 to 8.2 pigs per litter from 1930 to 1994 (Rothschild, 1996). These studies highlighted the importance of live born pigs, their birth weights, and the decrease in birth weight when LS increases (Johnson *et al.*, 1999d). Some of these results also illustrate the need to consider the effect of selection of LS on other traits in the long term (Estany *et al.*, 2002).

Given the difficulty of efficiently increasing LS through direct selection, MAS is considered to have potential benefits for improving selection for these traits with low heritability. Clearly, it is necessary to identify genetic markers associated with reproductive traits in order to implement MAS. Markers associated with reproductive traits have been identified through two complementary approaches. First, physiological candidate genes, which comprise genes with known roles in the trait of interest, are scanned for polymorphisms, and tested for associations with variation in the trait (Rothschild *et al.*, 1996; Short *et al.*, 1997; Rothschild *et al.*, 2000; Drogemuller *et al.*, 2001; Linville *et al.*, 2001; Jiang *et al.*, 2001; van Rens *et al.*, 2002; Gladney *et al.*, 2004; Buske *et al.*, 2005; Vallet *et al.*, 2005a; Campbell *et al.*, 2008; Zhou *et al.*, 2009; Lin *et al.*, 2009; Muñoz *et al.*, 2010; Fernandez-Rodriguez *et al.*, 2010; Fernandez-Rodriguez *et al.*, 2011; Lan *et al.*, 2011). Second, unbiased genome scans with anonymous DNA markers, such as microsatellites and more

recently with thousands of SNPs, have been used to identify QTL with effects on reproductive traits (Rathje *et al.*, 1997; Rohrer *et al.*, 1999; Wilkie *et al.*, 1999; Yasue *et al.*, 1999; Wada *et al.*, 2000; Rohrer, 2000; Cassady *et al.*, 2001; de Koning *et al.*, 2001; Hirooka *et al.*, 2001; King *et al.*, 2003; Lee *et al.*, 2003; Beeckmann *et al.*, 2003; Holl *et al.*, 2004; Rodriguez *et al.*, 2005; Sato *et al.*, 2006; Buske *et al.*, 2006a; Tribout *et al.*, 2008; Bidanel *et al.*, 2008; Li *et al.*, 2009; Ding *et al.*, 2009; Noguera *et al.*, 2009; Onteru *et al.*, 2011; Onteru *et al.*, 2012).

Most traits of economic importance in pigs are quantitative in nature with a continuous range of values between low and high performing animals. For several reproductive traits, such as LS and TN, the values within the range are discrete. The regions of the genome controlling quantitative traits are termed QTL. QTL can be identified by linkage analysis (or QTL mapping) in structured pedigrees or in Genome-Wide Association Studies (GWAS) which exploit population-wide linkage disequilibrium. The former approach is described here. Crosses between genetically and phenotypically diverse lines represent a powerful design for QTL mapping experiments (Rohrer *et al.*, 1994). Low-resolution genetic maps of the pig genome, essential to initial identification of QTL, have been published (Ellegren *et al.*, 1993; Rohrer *et al.*, 1994; Archibald *et al.*, 1995; Rohrer *et al.*, 1996; Marklund *et al.*, 1996).

In this study, three separate MS x LW cross populations developed at The Roslin Institute were analysed to identify QTL with effects on reproductive traits, including OR, TN, LS, TBA and PS. This is the first report of a complete genome scan for reproductive trait QTL for this population. An earlier report was limited to an analysis of SSC8 (King *et al.*, 2003). Although fine mapping analysis for SSC8 is described in Chapter 3, this chromosome is included in the genome-wide analyses described here.

After identification of a QTL, the ultimate goal is to identify the responsible gene itself, and the causative mutation for the phenotypic variation. However, although QTL with moderate effects on the trait of interest can be identified in QTL mapping studies, as described here, the resulting low-resolution QTL maps are not enough to

identify genes due to the large confidence intervals of the QTL. These large confidence intervals are likely to harbour many hundred genes. Although this issue of too many candidate genes remains a problem, the sequencing of the pig genome has dramatically increased the number of mapped pig genes. A partial draft pig genome sequence (Sscrofa9) was released in late 2009 and is accessible in the Ensembl, NCBI, and UCSC genome browsers. About 17,500 protein coding genes have been identified in the partial genome sequence (Ensembl Gene Build: http://www.ensembl.org/Sus_scrofa/Info/StatsTable?db=core). A more complete genome sequence (Sscrofa10.2) will be released and annotated in the next few months (Archibald *et al.*, 2010).

2.2. Materials and Methods

2.2.1. The population structure

Three separate MS x LW cross populations were developed at The Roslin Institute over a period of eight years. These populations were defined as QTL1, QTL2.1, and QTL2.2, where the last two populations had a small number of grandparental individuals in common. The founder grandparental animals were purebred MS and LW pigs. The purebred MS pigs were derived from an importation of 11 males and 21 females from the Jiadan county pedigree on the Lou Tang research farm in China in 1987 (Haley *et al.*, 1992). The first animals used in the present study were second-generation descendants of these imports. The purebred LW pigs were from a control population derived from a broad sample of LW genotypes present in the UK in 1982. Reproductive traits were recorded for 216 F_2 females.

All F_0 animals were unrelated (Walling *et al.*, 1998). The F_1 parents were produced through reciprocal crosses of F_0 purebred founder animals (MS male × LW female, and LW male × MS female). From the F_1 offspring, seven boars were mated to 25 sows of a different grandparental pairing, producing F_2 offspring in 43 full-sib families. Each F_1 sow had up to two litters of F_2 pigs. The resulting F_2 female offspring were mated to one of a few selected purebred LW boars, and various

reproductive traits were recorded. In total, the present study included 35 F_0 (13 males and 22 females), 94 F_1 (14 males and 80 females), and 216 F_2 (all females) individuals.

The trait-recorded F_2 animals had a minimum live weight of 85 kg at the start of each experiment, and they were reared indoors on standard commercial growth rations provided *ad libitum* until the time scheduled for first mating. Mating for each of the F_2 individuals in the two different year groups took place in two 6 - week periods. Gilts in the first age group (age group 1 - AG1) were 8 - 11 month of age, corresponding to the animal's first parity. They were then remated at 13 - 17 month of age. Individuals in this second age group (age group 2 - AG2) mostly had their second parity; a few who had an unsuccessful first mating had their first litter at this later age. All gilts and sows were observed daily for signs of oestrus and were mated on the same day as detection.

2.2.2. Phenotypic trait data

The phenotypic trait data had been recorded prior to the start of the present study, as the population was maintained on at The Roslin farm facilities from 1987 to 1996. The trait data were introduced into The Roslin ResSpecies database (www.resspecies.org). The ResSpecies database provides a secure and flexible environment for storing the data required for linkage and QTL analysis.

At 8 - 11 months of age the F_2 gilts were mated. At 5 - 20 days after mating, the weight of the animal and the number of CL on the left and right ovaries was recorded by laparoscopy and used as an estimate of OR. In addition, the TN on each side of each gilt and sow was counted. Some gilts and sows then returned to oestrus and, if they did so within the 6-week mating period, then they were remated. For those animals successfully remated in this manner, no record exists of the relevant number of CL, because the mating occurred after laparoscopy. These procedures were repeated for the same animals approximately 5 months later. This mating process

resulted in AG1 animals with parity 1 (8 - 11 months old) and AG2 with parity 1 or 2 (13 - 17 months old).

Trait information was retrieved from ResSpecies for QTL analysis. Before conducting QTL analysis, some traits needed to be calculated and others derived from those recorded. PS was calculated as the total number of piglets born or LS divided by OR for those animals in which their farrowing records corresponded to the OR recorded. It was assumed that the total number of CL reflected the maximum potential LS, and therefore the maximum value for PS was one. GL (in days) was calculated as the difference between the age of a gilt/sow at mating and its age at farrowing, with values around 114 days. Total OR and total TN were calculated from the partial numbers obtained from unilateral recording of each of the traits. The QTL or experimental group, in relation to the population to which the individual belonged, was added to the trait data for classification purposes. First group, QTL 1 (QTL group 1), was the oldest population and the other two, QTL 2.1 (QTL group 2) and 2.2 (QTL group 3), were subgroups from a bigger population from the same year. Once all the traits were calculated and the QTL group was added, the file was divided in two on the basis of age group. After, individuals with any missing measurements, with OR measurements which did not correspond to the LS data or with PS higher than one were removed from the file prior to analysis.

2.2.3. DNA samples

At the end of the experiment, the animals were slaughtered at around 18 - 22 months of age. DNA was prepared by standard procedures from spleen tissue, which had been collected *post-mortem* and stored at -70ºC. Briefly, a piece of frozen spleen sample, was cut with a sterile ostotome and placed in a mortar with liquid nitrogen in it and the sample was ground to a fine powder with a pestle. Once the liquid nitrogen in the mortar had evaporated, the powdered spleen was scraped using a scalpel blade into a 50 ml tube containing 6 ml of Nuclear lysis buffer pH 8.2 (10 mM Tris-HCl (Fisher Scientific Ltd., Loughborough, UK), 400 mM NaCl (Fisher Scientific Ltd.), 2 mM Disodium EDTA (Fisher Scientific Ltd.) pH 8.2, and distilled water).

Immediately, 200 µl of 20% SDS (20% w/v Sodium dodecyl sulphate, Amersham, Buckinghamshire, Biosciences) was added and mixed gently. Proteinase K (2 ml) (Roche, Welwyn Garden City, UK) of a stock solution (1 mg/ml) prepared in 1% SDS and 2 mM di-Sodium EDTA, was added and mixed well. The sample was left at 55ºC overnight in a shaking water bath. After this incubation, 2 ml of 6 M saturated NaCl was added and the tube was shaken vigorously for 15 seconds. The tube was centrifuged at 3,000 x *g* for 20 minutes in a bench top centrifuge (Eppendorf 5810R Bench top, Cambridge, UK) at 20ºC to avoid precipitation of SDS. The supernatant was decanted carefully into a clean 50 ml screw-top tube.

The DNA was precipitated in the 50 ml tube with an equal volume of isopropanol (Fisher Scientific Ltd.), the tube was inverted to mix the sample and the precipitated DNA was visible as white stringy fibres. The DNA was spooled with a sealed-end glass Pasteur pipette and washed by immersion in 5 ml of 70% ethanol in a 15 ml sterile polypropylene tube. The precipitated DNA attached to the glass pipette was removed from the wash and placed 'DNA end' up in a rack beside a 2 ml polypropylene screw-topped tube to air-dry for a few minutes. The sealed end of the pipette was placed into the 2 ml tube and re-suspended in 1 ml TE pH 8.0 (1 M Tris-HCl pH 8.0 and 0.5 M EDTA pH 8.0). After one hour the pipette was removed making sure the DNA remained in the tube. The DNA was left to dissolve at room temperature for around 12 hours inverting occasionally and then placed at 4ºC for a few days in order for the DNA to relax. Once relaxed, DNA concentration and quality were estimated on the Nanodrop ND-1000 (Labtech International Ltd., East Sussex, UK) and checked by electrophoresis on a 0.8% agarose gel (Sigma-Aldrich, Dorset, UK). Working dilutions at a final concentration of 12.5 ng/µl were prepared in 96-well plates (Thermo Scientific, Northumberland, UK) for all the samples and stored at 4ºC.

2.2.4. Genotyping of microsatellites markers

The genotypes of the F_2 trait-recorded females, their F_1 parents, and their purebred grandparents were determined for a total of 140 polymorphic genetic markers

previous to this study (Appendix 1). A total of 13 further markers were genotyped in another laboratory (Guelph, Canada) for this population. A total of a further 22 microsatellites located across several chromosomes were genotyped in this study. From the total of 175 markers distributed across the whole porcine genome, 158 were used in the initial analysis. In a second stage, after further genotyping the total of 174 markers (one marker was discarded as the data were unreliable and too sparse) provided coverage of 19 linkage groups. The information for the markers previously genotyped at The Roslin Institute or in other laboratories was available in the ResSpecies database. The information for the markers genotyped in this study was also loaded into the ResSpecies databases.

The optimal polymerase chain reaction (PCR) conditions for each specific primer pair were determined and PCR amplifications performed in 96-well PCR plates (Thermo Scientific) to determine the allelic size(s) for each microsatellite marker for a total of 307 samples. For each microsatellite marker one of the PCR primers was labelled with one of four different fluorescent-labels (FAM, VIC, NED, PET) (Applied Biosystem, Warrington, UK).

PCR reactions were performed in 10 µl reaction volumes containing 1 µl of 10x PCR Buffer + 15 mM Mg (Roche), 1 µl of 2 mM dNTPs (Thermo Scientific), 0.5 µl of each primer (10 pmol/µl), 0.05 µl of TAQ DNA polymerase (Roche) 5 U/µl, 4 µl of 12.5 ng/µl genomic DNA, and 2.95 µl of Milli-Q (Millipore Corporation, Watford, UK) water. Once the samples were placed in the corresponding well, the plate was sealed and placed in the thermocycler (MJ research PTC-225 Thermal Cycler, Peltier thermal cycler, Bio-Rad Laboratories Ltd., Hertfordshire, UK). The conditions for the PCR reaction were: 4 minutes of denaturation at 95ºC, 30 amplification cycles consisting of denaturation at 95ºC for 30 seconds, annealing at annealing temperature (Tm) (specific for each marker) for 30 seconds, and extension at 72ºC for 45 seconds; following these cycles, a further 10 minutes extension at 72ºC was performed.

After the first QTL analysis, some extra markers were genotyped for a number of chromosomes. At this stage a different source of Taq DNA polymerase was used by

the laboratory. PCR reactions were performed in 10 µl reaction volumes containing 1 µl of 10x PCR Buffer + 20 mM $MgCl_2$ (Roche), 1 µl of 2 mM dNTPs (Thermo Scientific), 0.5 µl of each primer (10 pmol/µl), 0.08 µl of FastStart Taq DNA polymerase (Roche) 5 U/µl, 4 µl of 12.5 ng/µl genomic DNA, and 2.92 µl of Milli-Q (Millipore Corporation) water. The conditions for these PCR reactions, using the same thermocycler as previously, were: 5 minutes of denaturation and Taq activation at 95°C, 30 amplification cycles consisting of denaturation at 95°C for 30 seconds, annealing at Tm for 30 seconds, and elongation at 72°C for 45 seconds, followed by a further 10 minutes final extension at 72°C.

For each sample/animal PCR products for multiple microsatellite markers were pooled. PCR products from microsatellite markers labelled with different fluorescent tags or with non-overlapping size ranges can be pooled. These PCR product pools were diluted in Milli-Q water in 96-well plates to give fluorescent peak heights in the range of 1,000 - 4,000 units, an optimal fluorescence intensity for PCR products of pooled microsatellite markers. From this dilution, 1 µl was mixed with 20 µl (per well) of Genescan 500LIZ size standard (Applied Biosystems) / Hi Di formamide (Applied Biosystems) mix. The stock mix for a 96-well plate (Thermo Scientific) was prepared by adding 4 µl of standard to 1 ml of Hi Di formamide. Once all the samples were loaded into the plate, it was sealed and pulse centrifuged prior to being loaded into an ABI 3730x1 96-capillary DNA analyser (Applied Biosystems). The samples (pooled PCR products) were subjected to capillary gel electrophoresis with each sample being loaded onto a separate capillary and the fluorescence intensities of the labelled PCR products and internal size markers captured automatically as they pass the detectors.

The results from the ABI 3730x1 96-capillary DNA analyser were examined with GeneMapper software (Applied Biosystems). The software allows electropherograms showing peaks of fluorescence intensity for the PCR fragments and internal size markers to be inspected, allelic fragments identified and sized for each sample. These results were checked individually, and where unclear or no results were detected, the analysis was repeated. Once the results for every sample were checked, the information for each sample results was exported in an Excel worksheet. This file

included sample file, sample name, marker, allele 1, allele 2, size 1, and size 2. When genotypes had been generated for at least 90% of the population for a marker of interest, the data were pre-submitted to ResSpecies, to check the data for possible inheritance errors. Where genotypes were inconsistent with the pedigree, indicating potential genotyping errors or errors in sample handling, the relevant data were re-examined in GeneMapper. Animals/samples, for which marker genotypes were inconsistent with the pedigree as recorded, were removed from the samples for further genotyping earlier based on the results for the first 20+ markers. Where the apparent genotyping errors could not be resolved by re-analysing the data in GeneMapper, the relevant sample(s) were genotyped again or the inconsistent genotype calls omitted from the database.

2.2.5. Linkage map construction

The information needed to build the linkage map was retrieved from ResSpecies using a tool for exporting data in the correct format for CRI-MAP (Green *et al.*, 1990) and MultiMap (Matise *et al.*, 1994) analyses. CRI-MAP is a computer program for likelihood-based map construction and MultiMap an expert system computer program. MultiMap automates the use of the CRI-MAP algorithm facilitating the construction task without the need for extensive user intervention. These programs were used to build framework and comprehensive linkage maps based on the recombination events in the QTL mapping pedigree. It might be expected that the marker order in these population-specific linkage maps would more often than not be the same as in reference linkage maps derived from other populations. However, as recombination frequencies along chromosomes can vary between families and individuals, the use of reference rather than population-specific linkage maps can result in errors in QTL mapping analyses.

SSC	Marker name	Primers	Primer sequence 5' - 3'	product size in bp	Tm
1	CH242-501j10	501j10-FAM	GGC TCA AGT ACC TGG ACT TAG TCT GC	224	60
		501j10-R	TGC CAC ATT TCT AGG CAC ACA GTT		
13	SW344	SW344-VIC	AGC TTC GTG TGT GCA GGA G	150-182	55
		SW344-R	GTA GTG GTC CAA AGA GAG TGC C		
13	SW2448	SW2448-VIC	CTC AGG GAC TTA TCC TCA GTG G	198-215	58
		SW2448-R	GAG GTG GGA TTT GGT CCA G		
13	SW1105	SW1105-VIC	TTC AAT TCA AAG AAG TGT TTG TG	105-139	60
		SW1105-R	GGT CGA TGA TGC TCA CAC C		
13	SW225	SW225-PET	AGG ACC CAC CAA GAG TTA CC	94-116	55
		SW225-R	TGC TGG TAA TGG GTG ATT AGG		
13	S0282	S0282-R	AAC TTC CAT ATG CCA CAG GTG C	112-142	60
		S0282-NED	AGT GGA ACA GAA TGG AGA GCC C		
18	SY4	SY4-FAM	TGT AAA AGA TTT AAT AGC CTG CCT C	144	62
		SY4-R	TGG TTT ATT CTT TCA TGA TTT CAT G		
18	SY31	SY31-FAM	TAG TAG CTG CAC ATG GTG TAA TTT	182	55
		SY31-R	TTG TGT AAA AAG GTA GAA AAC GC		
18	INHBA	INHBA-NED	CTC GTG TTC TCT TAC CAG AAG G	256	58
		INHBA-R	ACC CAG GTC GTA AGG TAT GTC		

Table 2.1 Microsatellite markers for fine mapping. The marker details presented include: location of the marker in the genome (SSC), name of the marker, fluorescent-tag for the labelled primer, primer sequences, size of the product to be amplified (bp), and the optimal annealing temperatura (Tm) for each pair of primers.

Both CRI-MAP and Multimap require the genotype and pedigree data in specific formats and require a defined set of files named in a specific manner. The file name format is chrN.xxx where N is an integer (conventionally the number of the chromosome of interest) and xxx is one of series of specific suffices analogous to **.doc**, **.xls** in Microsoft file names. Data exported from ResSpecies were saved into a text file, edited to remove control and formatting characters and renamed as a **.gen** file. The genotypes and the relationships between animal in the pedigree are presented in a specific format in the **.gen** file. Missing genotypes are inferred when possible from the parental genotypes by CRI-MAP. The next two files required for the analysis, **.dat** and **.loc**, were generated with the "**lispcri prepare**" function, specific for the use of MultiMap in the analysis. The **.loc** file contains a list of the marker loci and the number of informative meioses and phase known informative meioses for each marker. Once the three key files were generated, Emacs, a text editor, was used to create new files, **.ordj** and **.hash**. The former contains a single line: (()) and the latter contains a single line (nnnnn)() where nnnnn is the checksum for the **.dat** file. MultiMap requires a file (**.ordh**) listing the order in which the markers should be brought into the construction of the map. An appropriate strategy is to start with the most informative markers. In-house scripts have been developed to sort the marker list in the **.loc** file from the marker with the highest number of phase known informative meioses to the marker with the least phase known informative meioses and to write the ordered marker identifiers to an **.ordh** file. The makenames routine was used to generate a **.names** file containing the names of the marker loci from the **.loc** file. The **.names** file contains a list of all the markers in the order that they appear in the **.gen** and **.loc** files and then a second list of all the markers in same order but with the list enclosed in ( ). Editing the **.names** file to delete one or more marker names from the second list has the effect of excluding the deleted markers from the subsequent analysis. Thus, for a data set a variety of analyses using subsets of the markers can be explored without the need to create new **.gen** and all derived files.

The final file required was an **.input** file that consists of a script with all the commands and parameters for the construction of the linkage maps with MultiMap.

The parameters, which can be defined within the script, include minimum and maximum recombination distances between markers, the $\log_{10}$ likelihood odds for placement of markers with CRI-MAP's ALL routine, the $\log_{10}$ likelihood odds for analyzing an order with CRI-MAP's FLIPS routine, and lists of markers which can be haplotyped with the recombination interval between them set to zero.

A MultiMap session was initiated by invoking the following command: MultiMap. Once the session was opened, the analysis was run with ***multimap n***, where *n* corresponds to chr*n*.xxx in the filenames as described above. Initially, framework maps (Keats *et al.*, 1991) were constructed with an *odds-threshold 3.0*, or LOD ($\log_{10}$ odds) score of ≥3 (equivalent to odds of ≥1000:1) such that the reversal of the order of any pair of loci reduced the LOD score by more than 3. These framework maps were subsequently extended to generate comprehensive maps (Keats *et al.*, 1991) by inserting any non-framework markers for the chromosome of interest into their most likely positions by progressively reducing the odds-threshold.

The order and orientation of the linkage maps were investigated for consistency with published maps (http://www.marc.usda.gov/genome/swine/swine.html). The resulting linkage maps were also checked with the Chrompic option in CRI-MAP to identify putative double-recombinant events in short map distances (i.e. < 5cM). Such putative double recombination events are unlikely and highlight potential genotyping errors. The relevant genotypes results for the individual and marker together with the family results as well as results for markers surrounding it were re-examined. In the cases where both alleles had the same size and this result did not correspond with the inheritable alleles, GeneMapper results were checked for a possible weak peak for a second allele. If the correction was not possible that result was removed, and the analysis for the map was repeated. The final result was a map with the position of the markers in the chromosome in cM, the Kosambi distance between markers, and the recombination fraction for each marker (Theta). The linkage maps developed were then used for the QTL scan. After fine mapping of a chromosome region with extra markers, the map for that chromosome was built again repeating all the steps.

2.2.6. QTL scan analysis

The method used for QTL analysis of a three generation pedigree, derived from a cross between outbred lines, involved the use of regression-based interval mapping, and was effected using the GridQTL web interface enabling covariates and fixed effects to be fitted (Seaton *et al.*, 2002). A fixed QTL allele model, in which genetically distinct founder lines (MS and LW pigs in this case) were assumed to be fixed for alternative alleles at the QTL affecting the trait of interest, was used for the QTL scan analysis (Haley *et al.*, 1994).

The traits studied were PS, LS, TBA, OR and TN. Each reproductive trait measured for animals in AG1 was investigated individually for evidence of QTL in the genome. For all QTL analyses, experimental group (QTL1 (1), QTL2.1 (2), and QTL2.2 (3)) was included as a fixed effect, together with litter (1), sex (all animals were females) and parity (one for AG1 and 2 for AG2). As covariates, age at mating, weight at laparoscopy, age at farrowing, and GL were tested in the model, and the one(s) with more significant effect in the trait was used for the final analysis. The backcross-F_2 analysis tool was used for the one-QTL analyses and each chromosome scanned at 1 cM intervals for evidence of QTL with effects on the trait of interest.

Prior to the QTL analysis three different files were prepared: the genotype file with the allele size information for each individual for all the markers (exported from ResSpecies); the map file with details of the linkage maps (manually prepared with the results from 2.2.5 above, with the Kosambi distances); and the phenotype file with the trait information for all animals for the traits of interest. The phenotype file included the following traits: family number, litter, sex, parity, QTL group, age at mating, laparoscopy weight, age at farrowing, GL, QTL group, OR, TN, LS, TBA and PS, prepared as described in 2.2.2.

A genome-wide analysis with 1,000 permutations fitting the appropriate covariates was initially carried out to investigate the presence of significant QTL at chromosome and genome level. In these analyses, covariates were added and their

effects investigated. The F-values (variance ratio) were used to determine the significant threshold levels that were determine for chromosome-wide and genome-wide analysis using permutation and they were accepted when 5% significant threshold was reached (data presented). A bootstrap with resampling analysis was then carried out using 1,000 permutations of the trait data, to determine approximate confidence intervals for the QTL locations. In addition, the genetic background effect was investigated for all significant QTL found. For this analysis the position in cM of each QTL for a trait were fitted as genetic background effects together and then each of them was excluded, one at a time, in consecutive analyses. QTL reaching 5% and 1% chromosome-wide significant level are presented for each analyses. For example, if three QTL were found in SSC2, SSC5, and SSC11, in an initial analysis (1) all the QTL were fitted as background genetic effect. In a second analysis (2), SSC5 and SSC11 were fitted, in a third analysis (3) SSC2 and SSC11, and in the last analysis (4) SSC2 and SSC5 were fitted.

2.3. Results

2.3.1. Phenotype data

The phenotypic trait data were exported from ResSpecies. There was information for a total of 216 gilts/sows grouped in 77 families, from which 21 gilts were missing litter records. This resulted in a total of 195 animals left. From this, 167 had information for parity 1 at AG1 (8 - 11 months at mating). These gilts were mated for a second parity at AG2 but some of them were unsuccessful. The AG2 (13 - 17 months) had information for a total of 159 gilts and sows, from which 131 were the second parity of AG1 gilts, and 28 gilts had their first parity at this age. The data were divided into two files by age of the gilts/sows, AG1 (167 gilts) and AG2 (159 gilts/sows).

Traits recorded	Age Group 1 (n=137)			Age Group 2 (n=120)		
	Range	Mean (±SEM)	SD	Range	Mean (±SEM)	SD
Ovulation Rate	9-28	17.21 (0.30)	3.53	7-30	18.26 (0.35)	3.79
Teat number	12-18	14.93 (0.12)	1.37	11-18	15.11 (0.12)	1.32
Litter Size	2-22	12.12 (0.33)	3.85	1-22	12.70 (0.31)	3.45
Number Born Alive	1-17	10.96 (0.29)	3.42	1-21	11.77 (0.30)	3.27
Prenatal Survival	0.11-1	0.71 (0.02)	0.19	0.06-1	0.71 (0.02)	0.18
Covariates						
Age at mating (days)	248-357	302.41 (1.84)	21.50	402-559	492.01 (3.22)	35.32
Weight at laparoscopy (kg)	90-195	142.41 (1.82)	21.26	110-245	171.42 (2.53)	27.76
Age at farrowing (days)	362-469	416.54 (1.84)	21.50	517-673	606.33 (3.23)	35.41
Gestation Length (days)	108-119	114.13 (0.15)	1.74	111-118	114.33 (0.13)	1.44

Table 2.2 Phenotypic data summary table. The Table shows traits recorded indicating the ones used as covariates, range of values, mean and standard error of the mean (SEM), and standard deviation (SD) for each trait, separated by Age groups. n = number of animals.

SSC1

Marker	cM
SW1515	0.0
SW64	9.2
S0008	24.9
CGA	34.6
S0122	41.7
S0082	57.2
S0155	70.2
SW1301	116.2

SSC2

Marker	cM
SW2443	0.0
SW256	24.2
SW240	48.3
FSHB-2	58.0
SW1026	64.2
S0091	67.3
SW395	68.7
S0226	75.9
SW1695	82.6
S0378	95.0
SW1879	99.7
S0036	135.4

SSC3

Marker	cM
SW274	0.0
SW72	41.5
SW2527	58.3
SW902	69.1
FSHR-1	85.8
S0167	95.4
S0002	113.6
SW590	139.9

SSC4

Marker	cM
S0227	0.0
S0301	27.6
S0001	40.0
S0217	53.4
S0073	63.0
S0214	66.1
SW445	87.6
S0097	107.8

SSC5

Marker	cM
SW413	0.0
SWR453	40.8
DAGK	52.8
S0005	67.3
IGF1	95.5
SW1954	109.8
SW967	129.5

SSC8

Marker	cM
SW2410	0.0
HD-1	2.8
SW2611	3.9
SW905	22.0
QDPR-1	38.0
SLIT2	49.1
SW268	54.3
SW7	72.3
KIT	78.6
GNRHR	80.9
SULTE1	81.2
S0017	81.2
AREG	81.8
FGG	84.0
S0225	97.1
S0794	98.9
KS192	101.2
SW763	102.3
S0793	107.2
SW1551	114.1
SW790	116.7
SW61	119.7
S0782	123.5
SPP1	132.4
IBSP	132.6
S0792	135.6
SW1980	137.5
KS904	143.3
S0178	148.1

SSC6

Marker	cM
S0035	0.0
SW1057	45.0
S0220	75.4
SW122	83.8
SW316	88.3
SW71	93.6
S0031	97.3
S0228	102.1

SSC7

Marker	cM
S0025	0.0
SW2155	26.2
TNFB	46.3
BMP5	47.3
SW2019	47.9
DAXX	49.2
S0102	55.7
S0066	65.7
SW632	84.1
S0101	109.2
SW764	129.7

SSC9

Marker	cM
SW983	0.0
SW911	34.3
APOA1	58.6
SW1677	65.6
S0295	85.5
SW174	104.4
SW749	124.0

SSC10

Marker	cM
SW830	0.0
SW443	29.0
S0070	65.5
SW1041	71.6
SW951	95.3
SWR67	111.7

SSC11	
Marker	cM
S0385	0.0
SW1632	19.9
SW151	40.5
S0230	48.5
SW703	69.6

SSC12	
Marker	cM
SW2490	0.0
S0143	5.3
SW957	29.9
SW874	50.6
S0090	63.0

SSC13	
Marker	cM
SW1378	0.0
S0076	11.9
S0068	48.6
SW398	61.6
SW1056	72.7
SW769	88.3

SSC14	
Marker	cM
SW857	0.0
SW2496	15.3
SW295	35.8
SW210	41.0
S0007	53.0
SW761	67.9
SW1557	78.2
SW2515	97.2
SWC27	100.2

SSC15	
Marker	cM
S0355	0.0
S0148	12.4
SW964	29.5
S0149	38.5
SW936	59.6
SW1119	89.8

SSC16	
Marker	cM
SW742	0.0
SW403	17.6
S0026	37.8
SW1897	63.5

SSC17	
Marker	cM
SW335	0.0
S0296-2	27.5
S0359	55.5
SW2431	79.0

SSC18	
Marker	cM
SW2540	0.0
SW1984	21.5
SW787	23.6
SW1682	30.7
S0062	31.5
S0120	32.5
S0306	37.8

SSCX	
Marker	cM
SW2456	0.0
SW1943	32.1
S0218	66.8

Table 2.3 Linkage maps for the 18 porcine autosome chromosomes and X chromosome. Each map indicates chromosome number (SSC), names of the markers, and position of each marker in cM. Haplotyped markers are not shown in this table, full list of markers in Appendix 1.

The files were cleaned up as described in 2.2.2 removing any erroneous data; all the necessary trait calculations were completed and the files were cleaned of individuals with missing values, and the concordance between OR and LS were checked. After this cleaning, AG1 file had 137 gilts and AG2 file 120 gilts/sows (96 in the second parity and 24 in the first parity). The range, mean, and standard deviation (SD) of the phenotypic data in these files were calculated (Table 2.2).

2.3.2. 1st QTL analysis

2.3.2.1. Genotyping and linkage map

The number of individuals with genotype results was checked for all the markers in the ResSpecies database and the markers with an adequate coverage of the population (i.e. with genotypes for 296 - 306 individuals), were chosen for the linkage map construction (information in Appendix 1).

The genotyping results were exported from ResSpecies and the construction of the map was performed individually for each linkage group as described in 2.2.5 using 3.0 as lod-score threshold (Framework maps). The Chrompic results were checked and where errors were found the genotypes were masked in ResSpecies for each individual due to the impossibility of changing the genotypes previously measured.

These linkage maps contain between 3 and 12 markers per chromosome, with a total of 125 markers, excluding chromosome 8. Together with the 33 markers in chromosome 8, these maps cover 1902.4 cM (Table 2.4).

	Number of markers	cM
SSC1	8	116.2
SSC2	12	135.4
SSC3	8	139.9
SSC4	8	107.8
SSC5	7	129.5
SSC6	8	102.1
SSC7	11	129.7
SSC8	29	148.1
SSC9	7	124.0
SSC10	6	111.7
SSC11	5	69.6
SSC12	5	63.0
SSC13	6	88.3
SSC14	9	100.2
SSC15	6	89.8
SSC16	4	63.5
SSC17	4	79.0
SSC18	7	37.8
SSCX	3	66.8
Total	153*	1,902.4

Table 2.4 Linkage map summary. The chromosomes, number of markers per chromosome (SSC), and map coverage per chromosome in cMs are summarised. * The number of markers summarised here corresponds to the number of unique locations on the linkage map. The total number of markers genotyped is slightly larger, as markers between which no recombination was observed were treated as a haplotype.

2.3.2.2. QTL analysis

The analysis was performed as described in 2.2.6. First a genome-wide with 1,000 permutation analysis was run. Secondly, a bootstrap with resampling analysis was performed for each trait with 1,000 permutations for the whole genome. GL was used as a covariate for all the traits since it was the only covariate with effects on the traits under study, except for TN, where no covariate was used. The results of the analysis for the significant QTL are shown in Table 2.5.

Trait	SSC	Position (cM)	*F*-ratio	Estimated Effect		95% CI (cM) (start-end)	Significance threshold			
							Chromosome-wide		Genome-wide	
				Additive effect (±SE)	Dominance effect (±SE)		P <0.05	P< 0.01	P <0.05	P< 0.01
TBA	8	114	7.25	0.03 (0.39)	-2.17 (0.57)	0.5-140.5	5.71	8.29	8.53	10.48
LS	6	102	5.65	1.38 (0.45)	0.92 (0.65)	8.0-102.0	5.34	7.16	8.61	10.55
	8	114	6.12	0.07 (0.44)	-2.25 (0.64)	5.0-146.5	5.65	7.38		
	18	37	5.98	-0.49 (0.47)	-2.21 (0.65)	0.0-37.0	4.78	6.78		
PS	8	135	7.54	-0.03 (0.02)	-0.1 (0.03)	2.0-147.0	5.87	8.42	8.76	11.15
OR	7	56	7.45	-1.38 (0.45)	0.98 (0.58)	8.0-76.0	5.62	7.37	8.55	10.27
	13	38	8.45	-1.58 (0.48)	1.81 (0.81)	24.0-88.0	5.07	7.06		
	15	8	8.3	-1.82 (0.47)	1.06 (0.66)	2.0-60.0	4.87	6.63		
	18	37	5.16	-0.93 (0.42)	-1.48 (0.58)	0.0-37.0	4.58	6.79		
TN	5	57	10.28	-0.67 (0.15)	0.15 (0.24)	18.0-70.5	5.24	7.41	8.64	11.01
	18	0	4.82	-0.49 (0.15)	-0.15 (0.22)	0.0-37.0	4.56	6.48		

Table 2.5 Results from the genome-wide and bootstrap analysis. The table indicates the trait analysed, the chromosome (SSC) where a significant QTL was found associated with the trait, position of the QTL in cMs, F-ratio (variance ratio) for the QTL, estimated additive and dominance effects (± Standard error), confidence interval in cM, and chromosome-wide and genome-wide significant threshold for each QTL. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

The genome-wide permutation analysis revealed three QTL for OR (SSC7, SSC13, SSC15) and one for TN (SSC5) at 1% chromosome-wide significant level. This analysis also revealed one QTL for TBA (SSC8), three for LS (SSC6, SSC8, and SSC18), one for PS (SSC8), one for OR (SSC18) and one for TN (SSC18) at 5% chromosome-wide significance level. The QTL for TN on SSC5 was also significant at 5% genome-wide level. All the chromosome-wide significant QTL are shown in Figures 2.1 to 2.6. For each chromosome, not only the QTL plots for the trait with significant QTL are shown, but also the profiles for other related traits are shown for comparison purposes. In these Figures, the linkage map of the chromosome complete with marker names is shown on the x-axis and the statistical support for QTL at each position is shown on the y-axis.

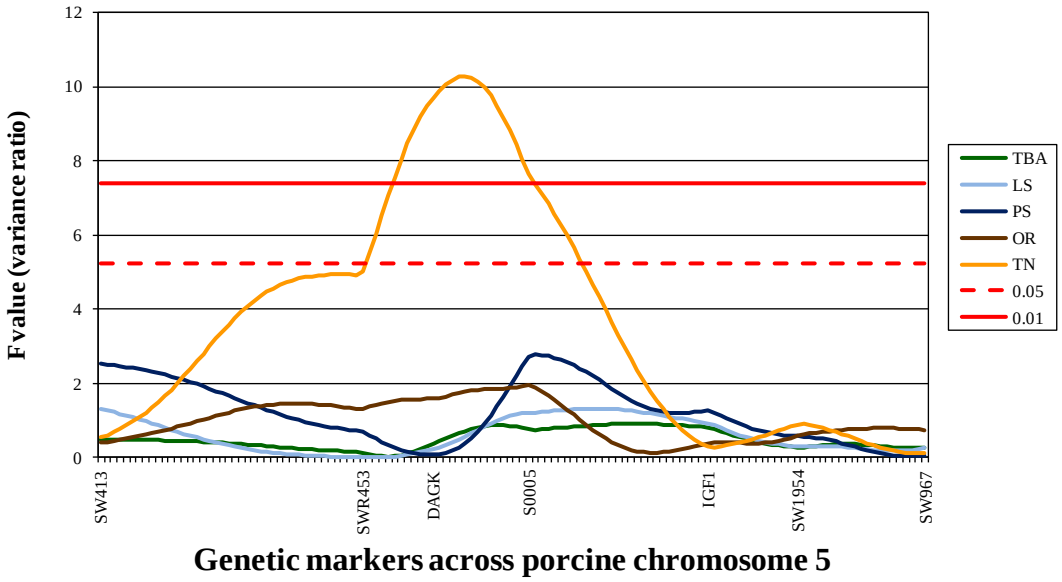

Figure 2.1 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC5. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

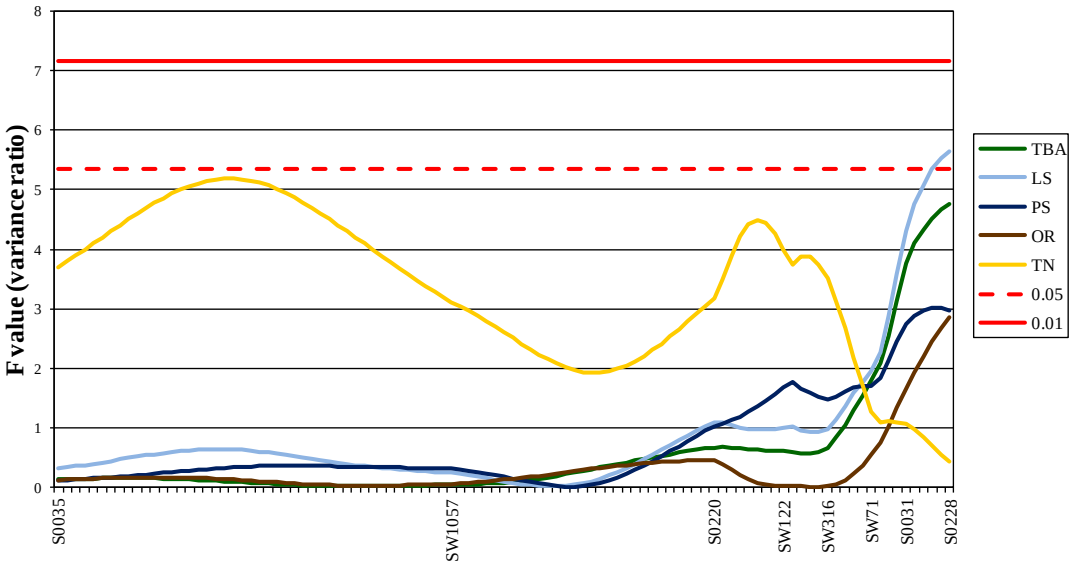

Figure 2.2 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC6. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

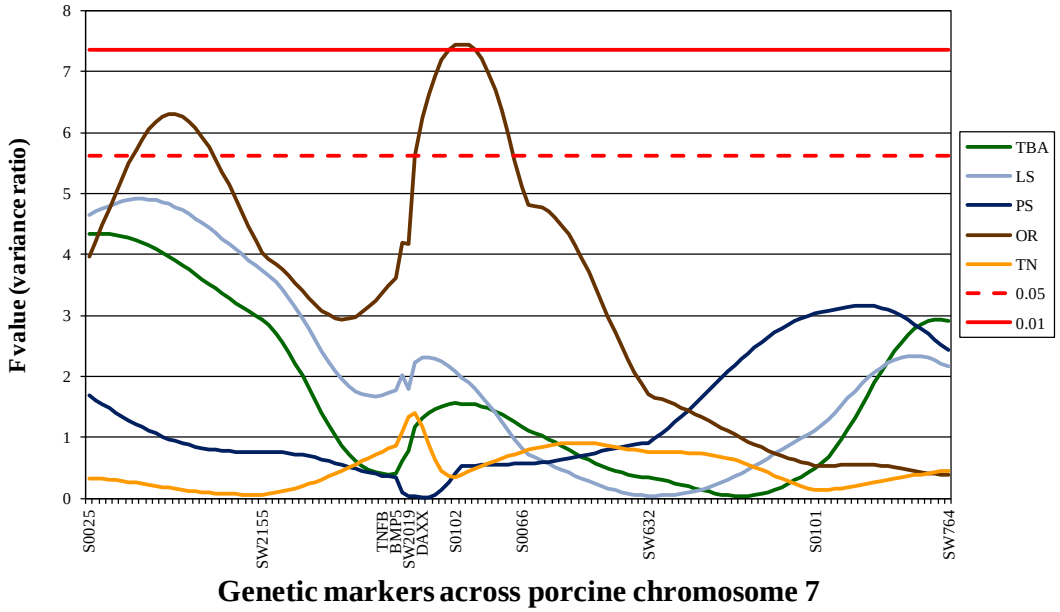

Figure 2.3 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC7. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

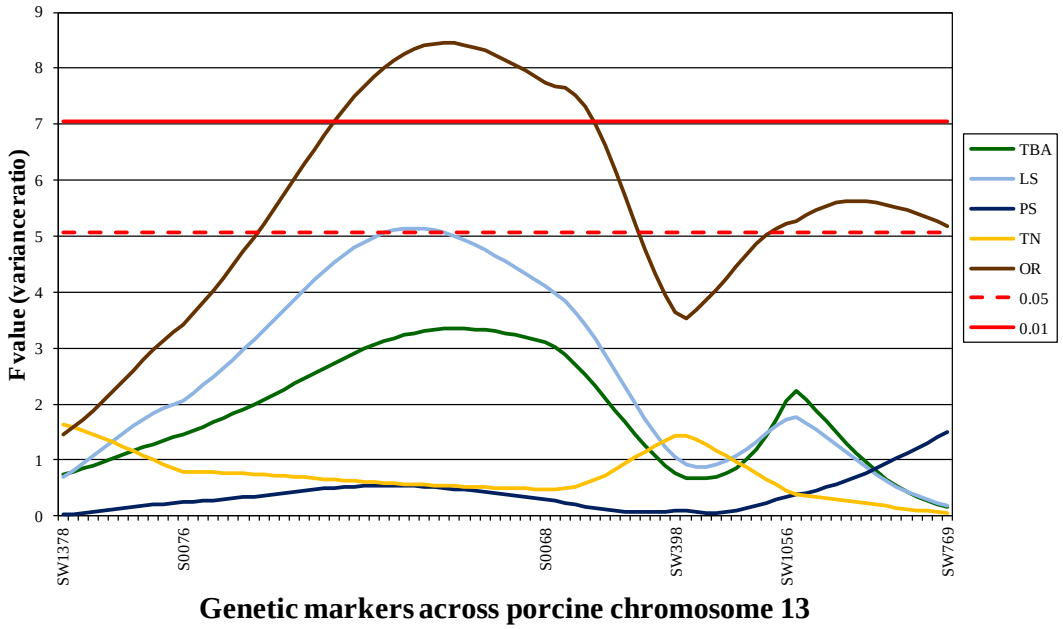

Figure 2.4 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC13. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

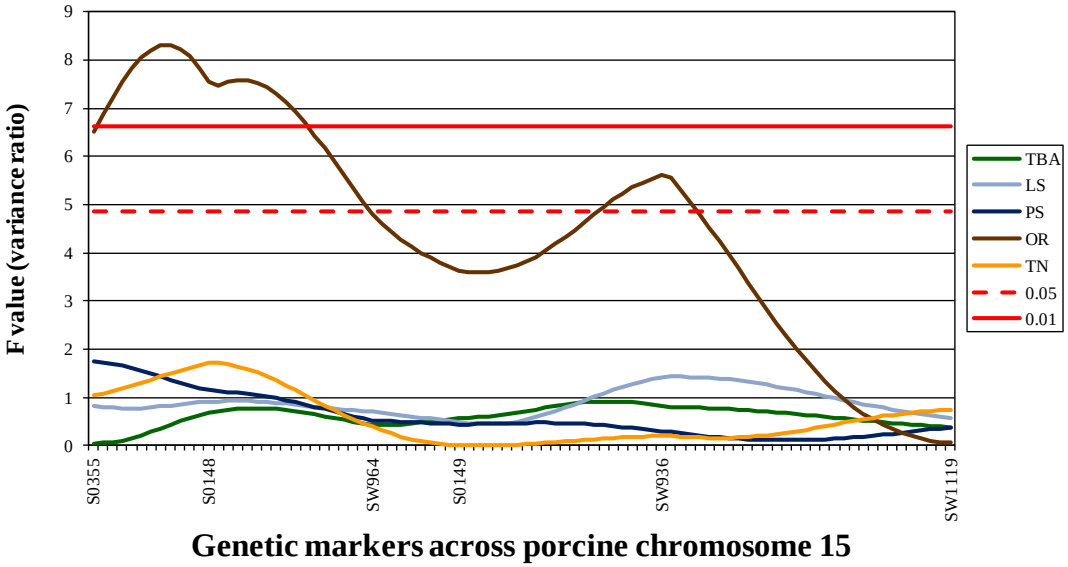

Figure 2.5 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC15. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

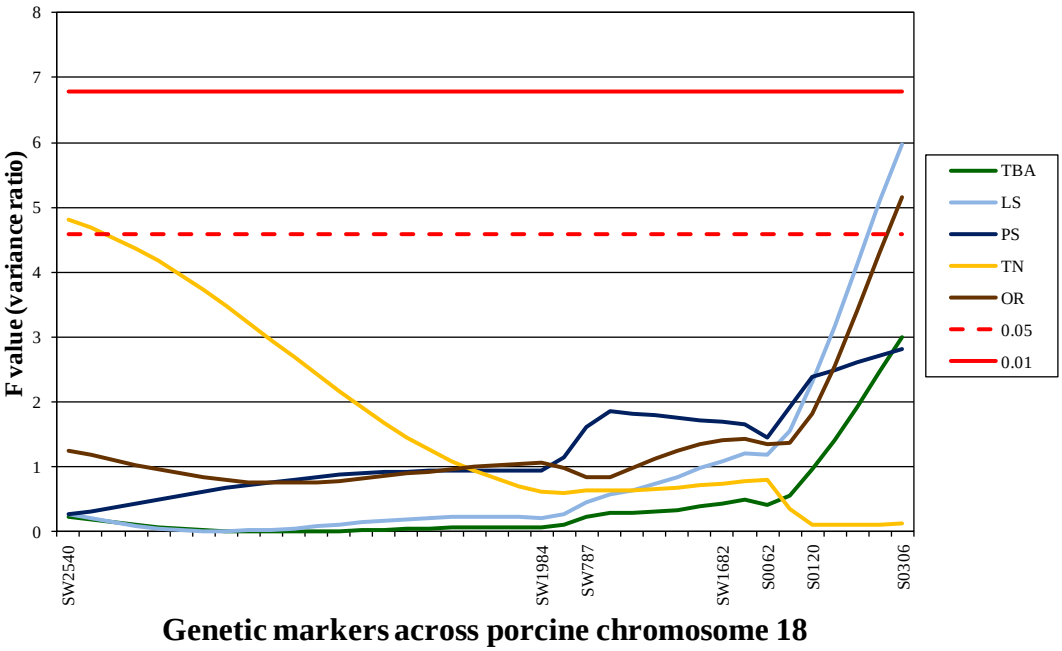

Figure 2.6 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC18. Chromosome-wide significance level at P<0.05 (Broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

The chromosome-wide significant QTL found in the genome-wide analysis were fitted as background effects and the analyses were repeated as described in 2.2.6, to investigate the genetic effect of these QTL and the interactions between them. The results of the background effect analyses are presented in Table 2.6. For the TBA trait, where only one QTL was observed (SSC8) in the first analysis, the background effect analysis revealed another putative QTL on SSC13 significant at 5% chromosome-wide level. For LS when all the putative QTL from the first round analysis were fitted, and for the other background analyses except when the SSC8 QTL was not fitted, a further putative QTL on SSC13 significant at the 5% chromosome-wide level was found in a similar position to the QTL for TBA described above. When only the putative QTL on SSC6 and SSC8 for LS were included as background effects, this SSC13 QTL was detected as well as the SSC18 QTL identified in the primary analysis.

Trait	Analysis	Input		Output				
		SSC	Position (cM)	SSC	Position (cM)	*F*-ratio	Significance threshold	
							P <0.05	P< 0.01
TBA	1	8	114	13	39	5.39	5.12	6.47
LS	1	6	102	13	35	6.94	5.07	7.33
		8	114					
		18	37					
	2	8	114	13	35	6.86	4.85	7.08
		18	37					
	3	6	102					
		18	37					
	4	6	102	13	35	7.87	5.09	6.93
		8	114	18	37	5.15	4.65	6.39
PS	1	8	114					
OR	1	7	56	14	21	7.54	5.30	7.36
		13	38					
		15	8					
		18	37					
	2	13	38	7	13	8.11	5.78	7.58
		15	8	14	20	7.35	5.32	7.01
		18	37					
	3	7	56	13	39	6.61	5.06	7.31
		15	8					
		18	37					
	4	7	56	14	21	7.48	5.43	7.99
		13	38					
		18	37					
	5	7	56	7	9	7.1	5.63	7.43
		13	38	14	20	7.68	5.42	8.05
		15	8					
TN	1	5	57	12	3	5.58	5.01	6.56
		18	0					
	2	18	0	5	56	9.1	5.47	7.65
	3	5	57					

Table 2.6 Result of the genetic background effects analysis. The Table shows traits, analysis number, input for analysis (Chromosome number (SSC) and position of QTL in cM), output or results of analysis (SSC, position in cM, F-ratio, and significant threshold at chromosome-wide level).

When the four putative OR QTL were included as background effects, and for the other analyses except when QTL on SSC13 was not fitted as a background effect, one new QTL appearred on SSC14. For this trait, only one of the initial QTL reappeared when it was not added as background effect, on SSC13. Only in some of the analyses with the QTL as background effect for OR, another QTL on SSC7 at 9-13 cM was shown. For TN there was a new QTL when both QTL were fitted as background effect (SSC12). One of the first QTL reappeared (SSC5) when it was not fitted as background effect. This QTL was the only one reaching 5% significant at the genome-wide level for the background analysis.

2.3.3. 2nd QTL analysis

2.3.3.1. Microsatellite genotyping and linkage map construction

A total of 9 additional microsatellite markers (Table 2.1) were genotyped to improve map coverage and resolution for three different chromosomes, SSC1, SSC13, and SSC18, as described in 2.2.4. The results are presented in Table 2.7. The construction of the comprehensive maps was repeated for all the chromosomes as described in 2.2.5, adjusting the odd threshold level to maximise the number of markers in the map without compromising the quality of the map. This extension of the framework maps to establish comprehensive linkage maps was not performed in the analyses described in 0. The results and a summary of the results are presented in Table 2.8 and Table 2.9, respectively. These maps hold a total of 174 markers covering 1901.5 cM.

SSC	Primers	Animals with genotypes
1	501J10	291
13	SW344	300
13	SW2448	301
13	SW1105	297
13	SW225	304
13	S0282	305
18	SY4	302
18	SY31	295
18	INHBA	304

Table 2.7 Number of animals genotyped for extra markers.

SSC1

Marker	cM
SW1515	0.0
501j10	0.9
SW64	9.2
S0008	24.6
CGA	32.9
S0122	38.0
S0082	50.9
S0155	60.1
SW1301	100.5

SSC2

Marker	cM
SW2443	0.0
SW256	24.2
SW240	48.3
FSHB-2	58.0
SW1026	64.2
S0091	67.3
SW395	68.7
S0226	75.9
SW1695	82.6
S0378	95.0
SW1879	99.7
S0036	135.4

SSC3

Marker	cM
SW274	0.0
SW72	41.5
SW2527	58.3
SW902	69.2
FSHR-H	85.8
S0167	94.8
S0002	113.1
SW590	139.4

SSC4

Marker	cM
S0227	0.0
S0301	28.0
S0001	40.6
S0023	53.6
S0217	54.4
S0073	63.6
S0214	66.7
SW445	88.2
S0097	108.4

SSC5

Marker	cM
SW413	0.0
SWR453	40.6
GDF11	52.3
DAGK	52.8
S0005	67.3
IGF1	95.4
SW1954	109.7
SW967	129.4

SSC8

Marker	cM
SW2410	0.0
HD-1	2.5
SW2611	3.7
SW905	19.3
QDPR-1	32.6
SLIT2	44.1
SW7	65.2
KIT	70.7
GNRHR	72.8
SULTE1	73.1
S0017	73.1
AREG	73.7
FGG	75.6
S0225	89.1
S0794	90.9
KS192	93.5
SW763	94.1
S0793	98.7
238o22b	101.3
27o17	103.0
SW1551	105.0
SW790	106.1
SW61	108.8
S0782	112.3
SPP1	121.2
IBSP	121.4
S0792	124.4
SW1980	125.9
443f10	130.9
KS904	131.3
S0178	136.2

SSC6

Marker	cM
S0035	0.0
SW1057	45.0
S0220	75.4
SW122	83.8
SW316	88.3
SW71	93.6
S0031	97.3
S0228	102.1

SSC7

Marker	cM
S0025	0.0
SW2155	26.2
TNFB	46.3
BMP5	47.3
SW2019	47.9
DAXX	49.2
S0102	55.7
S0066	65.7
SW632	84.1
S0101	109.2
SW764	129.7

SSC9

Marker	cM
SW983	0.0
SW911	34.3
APOA1	58.6
SW1677	65.6
S0295	85.5
SW174	104.3
SW749	124.0

SSC10

Marker	cM
SW830	0.0
SW443	29.0
S0070	65.5
SW1041	71.6
SW951	95.3
SWR67	111.7

SSC11	
Marker	cM
S0385	0.0
SW1632	19.9
SW151	40.5
S0230	48.5
SW703	69.6

SSC12	
Marker	cM
SW2490	0.0
S0143	5.3
SW957	29.9
SW874	50.6
S0090	63.0

SSC13	
Marker	cM
S0282	0.0
SW1378	14.1
S0076	25.0
SW344	35.3
SW2448	50.4
SW1105	56.4
S0068	57.8
SW225	63.9
SW398	70.5
SW1056	81.2
SW769	97.0
S0215	97.8

SSC14	
Marker	cM
SW857	0.0
SW2496	15.3
SW295	35.8
SW210	41.0
S0007	53.0
SW761	67.9
SW1557	78.2
SW2515	97.2
SWC27	100.2

SSC15	
Marker	cM
S0355	0.0
S0148	12.4
SW964	29.5
S0149	38.5
SW936	59.6
SW1119	89.8

SSC16	
Marker	cM
SW742	0.0
SW403	17.6
S0026	37.8
SW1897	63.5

SSC17	
Marker	cM
SW335	0.0
S0296-2	27.5
S0359	55.5
SW2431	79.0

SSC18	
Marker	cM
SW1808	0.0
SW2540	1.3
SY4	5.4
SW1984	22.6
SW787	24.6
SW1682	31.5
S0062	32.3
S0120	33.2
S0306	37.7
SY31	52.4
INHBA	54.7

SSCX	
Marker	cM
SW2456	0.0
SW1943	32.1
S0218	66.8

Table 2.8 Linkage maps for the 18 porcine autosomal chromosomes and X chromosome incorporating additional markers. Each map indicates name of chromosome (SSC), names of the markers, and position of each marker in cM. Haplotyped markers are not shown in this table, full list of markers in Appendix 1.

	Number of markers	cM
SSC1	9	100.5
SSC2	12	135.4
SSC3	8	139.9
SSC4	9	108.4
SSC5	8	129.4
SSC6	8	102.1
SSC7	11	129.7
SSC8	31	136.0
SSC9	7	124.0
SSC10	6	111.7
SSC11	5	69.6
SSC12	5	63.0
SSC13	12	97.8
SSC14	9	100.2
SSC15	6	89.8
SSC16	4	63.5
SSC17	4	79.0
SSC18	11	54.7
SSCX	3	66.8
Total	168*	1,901.5

Table 2.9 Summary of linkage map construction after fine mapping. The Table indicates chromosome number (SSC), number of markers per chromosome and cM cover by the markers. Map results different to the previous ones are underlined. * The number of markers summarised here correspond to the number of unique locations on the linkage map. The total number of markers genotyped is slightly larger, as markers between which no recombination was observed were treated as a haplotype.

2.3.3.2. QTL analysis

The initial scan for QTL with effects on the traits of interest was performed with the new linkage maps as described in 2.2.6. The results of these analyses are shown in Table 2.10.

Trait	SSC	Position (cM)	*F*-ratio	Estimate Effect		95% CI (cM) (start-end)	Significance threshold			
				Additive effect (±SE)	Dominance effect (±SE)		Chromosome-wide		Genome-wide	
							P <0.05	P< 0.01	P <0.05	P< 0.01
TBA	8	105	6.98	-0.03 (0.38)	-2.12 (0.56)	0.0-133.0	5.92	7.54	8.72	10.67
	18	49	6.02	-0.18 (0.46)	-2.37 (0.68)	5.0-53.0	4.62	7.00		
LS	6	102	5.65	1.38 (0.45)	0.92 (0.65)	7.0-102.0	4.85	6.66	8.96	11.21
	8	105	5.86	-0.03 (0.43)	-2.18 (0.63)	4.0-135.0	5.85	7.65		
	18	47	7.41	-0.36 (0.52)	-2.95 (0.77)	6.0-53.0	5.06	7.17		
PS	8	124	7.53	-0.03 (0.02)	-0.1 (0.03)	2.0-136.0	6.45	7.99	8.5	10.61
OR	7	56	7.45	-1.38 (0.45)	0.98 (0.58)	8.0-75.0	5.4	8.15	8.66	10.99
	13	56	8.42	-1.51 (0.4)	0.84 (0.56)	27.0-97.0	5.15	7.57		
	15	8	8.3	-1.82 (0.47)	1.06 (0.66)	2.0-60.0	5.33	6.96		
	18	42	5.28	-1.064(0.45)	-1.59 (0.66)	1.0-52.0	5.22	7.26		
TN	5	52	10.32	-0.63 (0.14)	0.12 (0.22)	17.5-69.0	5.52	7.47	8.93	11.11
	6	20	5.19	-0.75 (0.23)	-0.33 (0.26)	0.0-97.0	5.1	6.9		
	18	0	6.44	-0.55 (0.15)	-0.17 (0.21)	0.0-50.5	5.18	7.34		

Table 2.10 Results from the genome-wide and bootstrap analysis after fine mapping. The Table indicates the trait analysed, chromosome where a significant QTL was found associated with the trait, position of the QTL in cMs, F-ratio for the QTL, estimated additive and dominance effect (± Standard error), confidence interval in cM, and chromosome-wide and genome-wide significant threshold for each QTL. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

The genome-wide permutation analysis revealed putative QTL as follows: one QTL for LS (SSC18), two for OR (SSC13, SSC15) and one for TN (SSC5) at 1% chromosome-wide significant level. This analysis also revealed putative QTL at the 5% chromosome-wide significance level as follows: two QTL for TBA (SSC8, SSC18), two for LS (SSC6, SSC8), one for PS (SSC8), two for OR (SSC7, SSC18), and two for TN (SSC6, SSC18). The putative QTL for TN on SSC5 was also significant at 5% genome-wide level. The chromosome-wide significance threshold was defined by random permutations, thus it changed between analyses revealing a putative new QTL on SSC6 with effects on TN although no new SSC6 markers had been added to the map for these analyses. The graphics for these QTL are represented in Figures 2.7 to 2.11. As before, for each chromosome not only are the QTL plots for the trait with significant QTL shown, but also the profiles for other related traits are shown for comparison purposes. In these Figures the linkage map of the chromosome complete with marker names is shown on the x-axis and the statistical support for QTL at each position is shown on the y-axis.

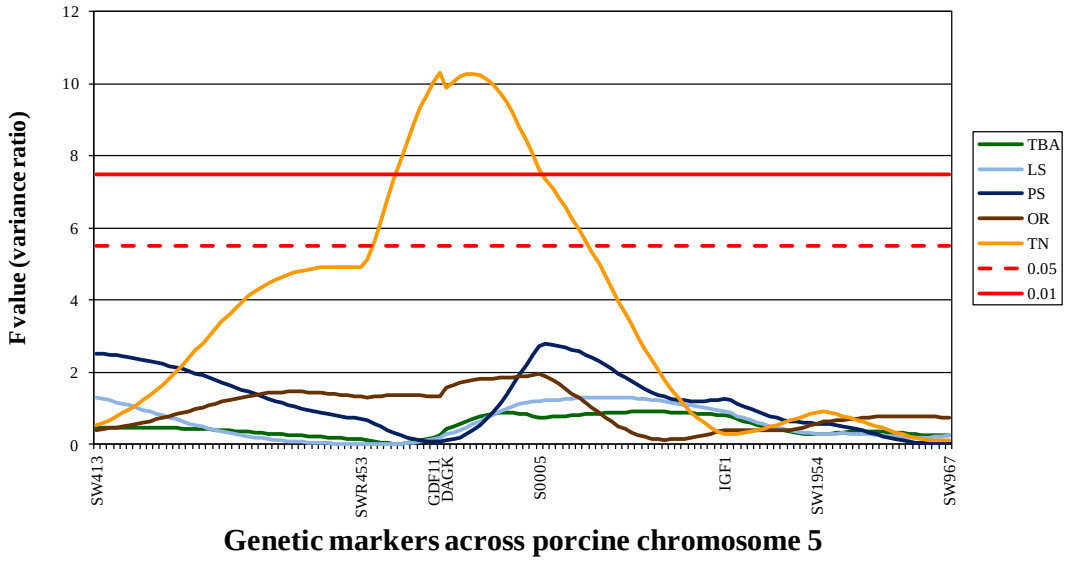

Figure 2.7 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC5 after adding another marker to the map (cf. Figure 2.1). Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

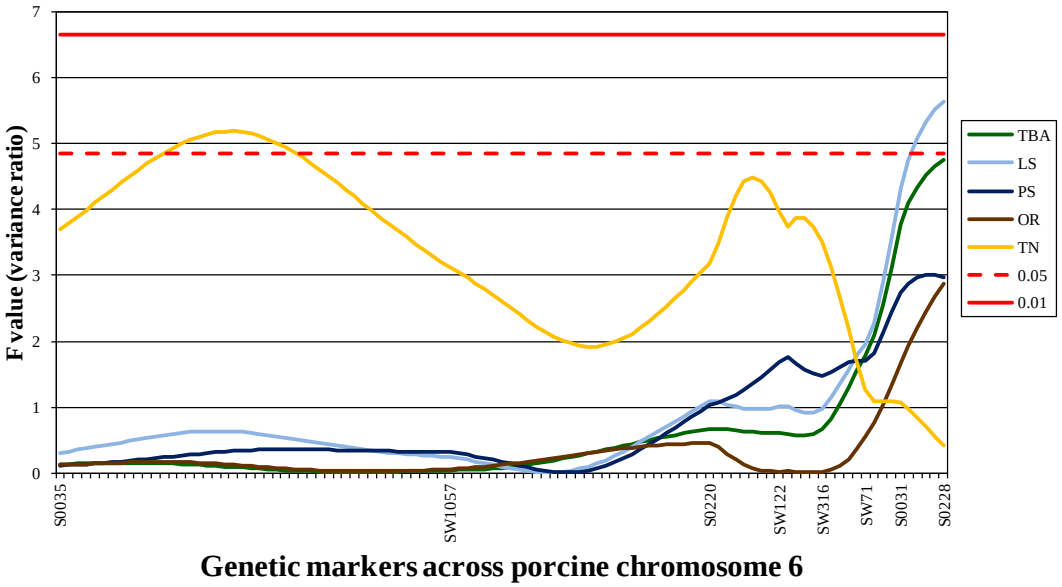

Figure 2.8 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC6 (cf. Figure 2.2**)**. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

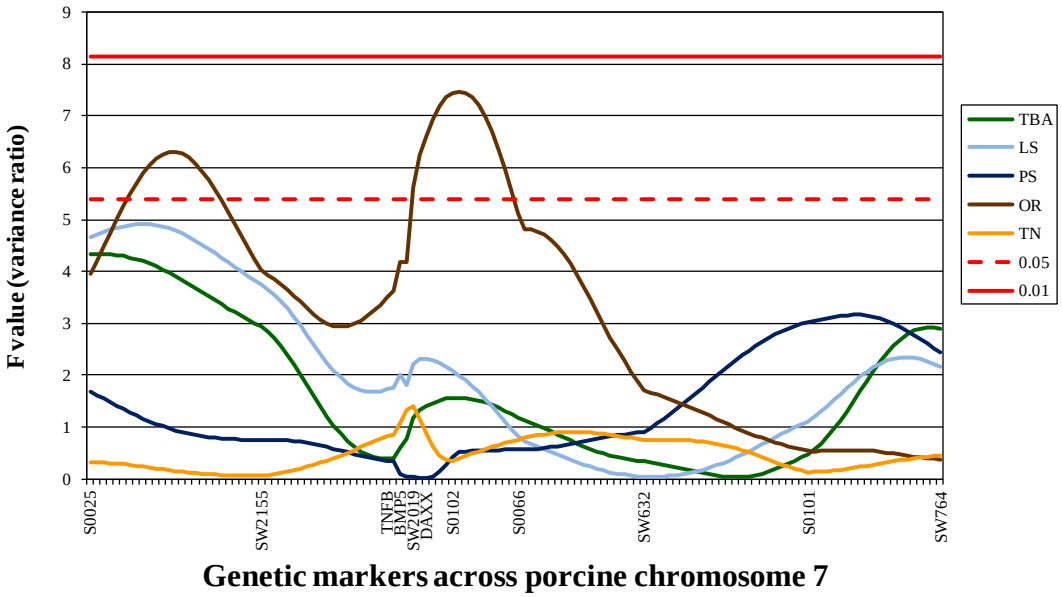

Figure 2.9 Interval mapping for QTL with effects TBA, LS, PS, OR and TN on SSC7 (cf. Figure 2.3). Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

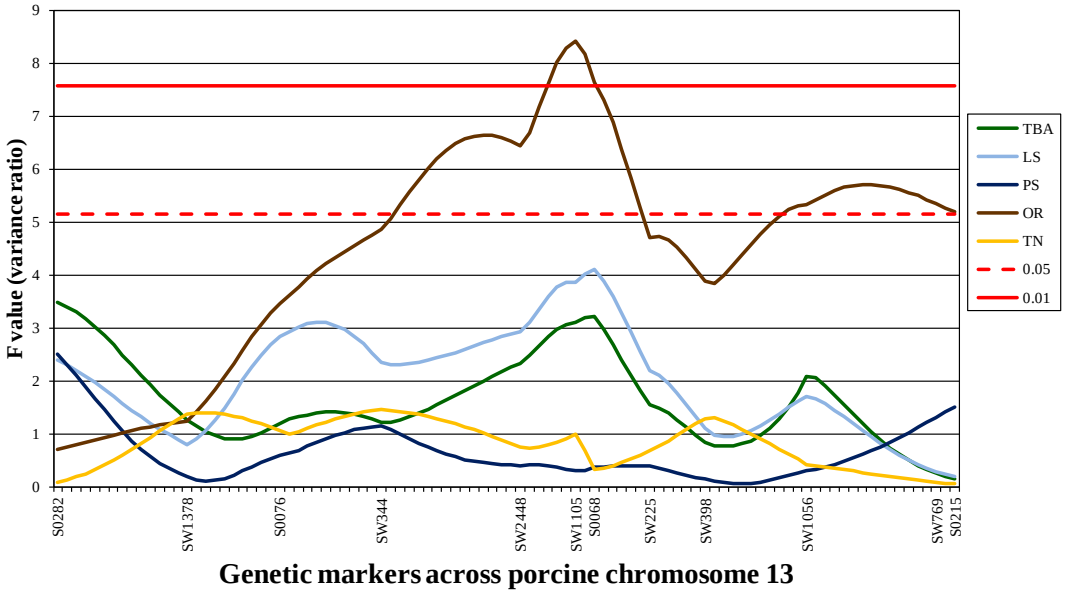

Figure 2.10 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC13 after fine mapping (cf. Figure 2.4). Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

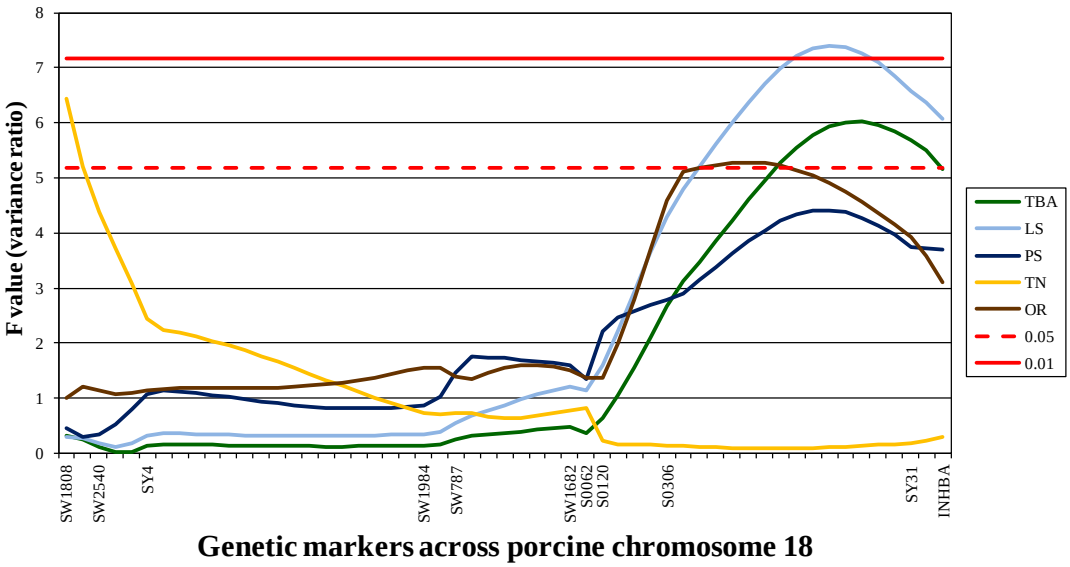

Figure 2.11 Interval mapping for QTL with effects on TBA, LS, PS, OR and TN on SSC18 after fine mapping (cf. Figure 2.6). Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

Trait	Analysis	Input		Output				
		SSC	Position (cM)	SSC	Position (cM)	*F*-ratio	Significance threshold	
							P <0.05	P< 0.01
TBA	1	8	105	13	58	5.51	5.34	7.24
		18	49					
	2	18	49					
	3	8	105					
LS	1	6	102	13	58	6.4	5.19	7.70
		8	105					
		18	47					
	2	8	105	13	58	6.57	5.77	7.82
		18	47					
	3	6	102					
		18	47					
	4	6	102	4	88	5.31	5	6.64
		8	105	13	58	6.13	5.47	7.47
PS	1	8	124					
OR	1	7	56	14	20	6.69	5.45	7.69
		13	56					
		15	8					
		18	42					
	2	13	56	7	12	7.29	5.36	7.08
		15	8	14	20	6.74	5.52	7.51
		18	42					
	3	7	56	13	55	7.62	5.45	7.48
		15	8					
		18	42					
	4	7	56	14	20	6.54	5.35	7.22
		13	56					
		18	42					
	5	7	56	7	9	6.8	5.36	7.16
		13	56	14	19	6.63	5.5	7.67
		15	8	18	48	5.64	4.89	6.41
TN	1	5	52	12	3	8.78	4.76	6.85
		6	20					
		18	0					
	2	6	20	5	56	8.52	5.25	6.92
		18	0	12	2	6.61	4.74	6.96
	3	5	52	12	3	5.88	4.90	6.75
		18	0					
	4	5	52	11	0	6.03	4.97	7.56
		6	20	12	3	6.67	4.59	6.18

Table 2.11 Result of the genetic background effects analysis after second analysis. The Table shows traits, analysis number, and input for the analysis (Chromosome number and position of the QTL), output or results of the analysis including chromosome number, position of the QTL, F-ratio and significant threshold at chromosome-wide level. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

The results of the background effect analyses are presented in Table 2.11. For the TBA trait, where two putative QTL were observed, the background effect analysis with both QTL revealed another putative QTL on SSC13, significant at 5% chromosome-wide level as observed previously. This SSC13 QTL was also observed in the background analysis for LS when the SSC8 QTL was included as a background genetic effect. When the putative QTL on SSC18 was excluded in the background analysis for LS, as a result of the analysis a new QTL on SSC4 appeared. For OR in all the analyses, except when the QTL on SSC13 was not fitted as a background effect, one new putative QTL appeared on SSC14. When SSC13 or SSC18 QTL were not fitted as background effects, separately, they reappeared. In the analysis where SSC7 or SSC18 QTL were not added as a background effect, respectively, a QTL in SSC7 at 9-13 cM was observed. For TN, where three QTL were found, a new putative QTL was found in all the background effect analyses (SSC12). The QTL on SSC5 reappeared when it was not fitted as a background effect. This QTL was the only one reaching 5% significance at the genome-wide level. Also, when SSC18 QTL was excluded, a further putative QTL on SSC11 appeared.

In the first analysis, a putative QTL with effects on LS and OR was detected on SSC18 in a region which harbours a potential candidate gene (*INHBA*) for reproductive traits. Therefore, additional SSC18 markers, including a marker within the *INHBA* locus were genotyped in the population and the QTL analyses repeated. After repeating the analysis, which revealed putative QTL for TBA, LS and OR on this chromosome region as shown in Table 2.100, the marker for this gene was fitted as a background effect in order to examine the possibility that variation in the *INHBA*

gene was responsible for the observed QTL effects. As a result of this analysis, no new QTL were found and the QTL at SSC18 did not reappear (results no presented).

2.4. Discussion

In this study, the whole genome of The Roslin LW x MS population was scanned for the first time for QTL with effects on reproductive traits (LS, PS, OR, TN, TBA). Previous analysis of this population for these traits explored only SSC8 (King *et al.*, 2003). A linkage-based approach for QTL detection was used exploiting three-generation F_2 intercross pedigrees in which the founder generation (F_0) were LW and MS pigs, not only increasing the number of chromosomes analysed compared to the previous analysis but also using additional markers to the ones used previously. These breeds exhibit significant differences in female reproductive performance and the QTL analyses were based on the assumption that the founders are fixed for different alleles at the QTL.

Similar linkage-based genome scans for QTL with effects on reproductive performance have been reported by others. The linkage maps in this study, comprising 174 markers in 19 linkage groups with a total map length of 1901.5 cM, were consistent with previous studies (Rohrer *et al.*, 1996; Rathje *et al.*, 1997; Bidanel *et al.*, 2001; Rodriguez *et al.*, 2005; Noguera *et al.*, 2009; Vingborg *et al.*, 2009). The number of the markers varies in these studies from 55 markers in 16 chromosomes (Rathje *et al.*, 1997) to 1042 loci in 19 chromosomes (Rohrer *et al.*, 1996). The linkage map coverage also varies in these studies that cover from 1364.3 cM in 16 chromosomes (Rathje *et al.*, 1997) to 2565 cM in 19 chromosomes with 136 markers (Bidanel *et al.*, 2001).

Analysis of The Roslin LW x MS populations revealed ten putative QTL on six different chromosomes with effects on four different traits, excluding the SSC8 QTL, results that are discussed in the next chapter. As a result of this study, a QTL for TN on SSC5 at 5 % genome-wide significant level was found, together with one for TBA, one for LS, two for OR and two for TN at 5% chromosome-wide significance

level, and one for LS and two for OR at 1% chromosome-wide significance level. These results were examined and compared with previous studies. Results from studies to identify genes with effects on reproductive traits in pigs using genome scans, physiological candidate genes, and functional genomics approaches have been reviewed recently (Onteru *et al.*, 2009). The pig QTL database (pigQTLdb) (http://www.animalgenome.org/cgi-bin/QTLdb/SS/index) also provides a valuable resource for comparing results from different QTL studies.

Three putative QTL with effects on TN were detected in this study, including the only QTL significant at the 5% genome-wide level. This latter TN QTL on SSC5, was mapped to a similar position of a QTL described in a previous study by Ding *et al.* (2009) in a White Duroc x Erhualian population, a commercial composite line and a Chinese indigenous breed with superb fertility performance, respectively. Two other studies, both of which exploited MS-European intercross F_2 populations, also reported TN QTL with locations which overlap with the QTL observed here on SSC5 (Lee *et al.*, 2003; Rodriguez *et al.*, 2005). Martinez-Giner and colleagues (2011) examined the gene encoding parathyroid hormone-like hormone (*PTHLH*), which has a role in mammary development, and which maps to SSC5, as a candidate gene for this TN QTL. From studies of PTHLH gene expression and an association study of a *PTHLH* polymorphism in the Iberian-MS population described by Rodriguez *et al.* (2005), they concluded that *PTHLH* was unlikely to be the gene responsible for the TN QTL effects. The SSC5 TN QTL detected in The Roslin population maps to the region that contains QTL with effects on ear size and erectness, for which there is very compelling statistical support (Wei *et al.*, 2007). Both traits – TN and ear size/erectness – could be viewed as body patterning traits which might be shaped by early developmental processes. Therefore, although the resolution with which these QTL can be mapped is limited, it is possible that a gene with roles in development could have effects on both traits.

The putative TN QTL located around the SW1057 marker on chromosome 6 has been observed in a earlier study of a population derived from the same founder animals as the population described here (Guo *et al.*, 2008). The putative QTL for TN detected on SSC18 is not corroborated by results from other studies. For a

multiparous animal such as the pig, TN is an important trait and can affect the ability of a sow to nuture her offspring. Variation in TN between individuals is evident from the phenotypic data (see Table 2.2). The QTL with effects on TN detected in this and other studies demonstrate that genetic contributions to TN variation exist (Figure 2.12).

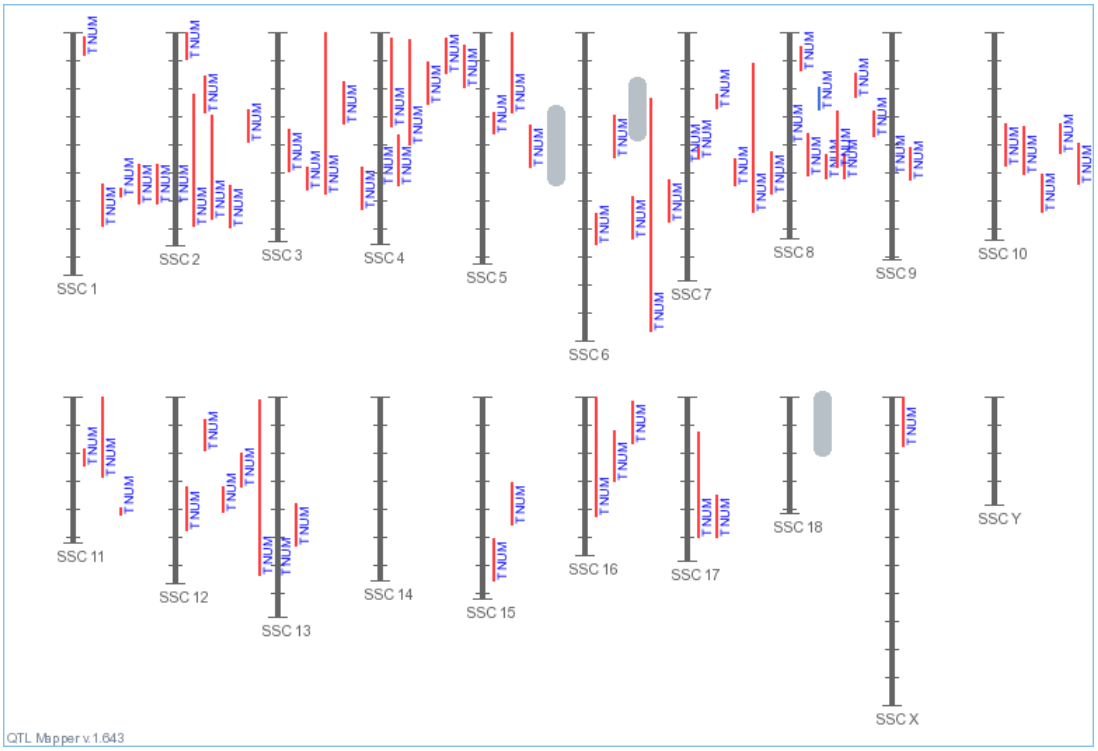

Figure 2.12 Graph representation of the QTL published for TN from pigQTLdb together with the QTL mapped in the present study for this trait. The QTL detected in the current study are shown as thick grey bars.

Four putative QTL with effects on OR were detected in this study, two of them at the 1% chromosome-wide significance level (SSC13, SSC15) and two at the 5% chromosome-wide significance level (SSC7, SSC18). Recently, Bidanel *et al.* (2008) found a QTL on SSC7 for OR at the 5% genome-wide significance level in a population similar to the one used in this study, a LW x MS F_2 population, and the position of this QTL was close to the one mapped in this study, with a positive additive effect. Also Wilkie *et al.* (1999) mapped a QTL for this trait in a MS x

Yorkshire population to this chromosome but in a different region than the one in the present study.

Bidanel *et al.* (2008) also mapped a QTL for this trait to SSC13, close to the one mapped in the present study in this chromosome. Rathje *et al.* (1997) mapped QTL to SSC13 and SSC15 in a LW x LR population. Although QTL on both chromosomes were mapped in this study, both mapped to different positions compared to the QTL reported by Rathje *et al.* (1997). This difference is possibly due to the different composition of the population. On SSC15 another two QTL for OR have been mapped in previous studies, one by Rohrer *et al.* (1999) in a White Composite x MS population in a position close to the QTL mapped by Rathje *et al.* (1997), and another one by Wilkie *et al.* (1999) in a MS x Yorkshire population. However, these SSC15 OR QTL reported by others, map to significantly different locations to those observed in the present study. As for TN, a QTL on SSC18 was detected for OR, but in another position, where no previous QTL were mapped.

For TBA, a QTL significant at the 5% chromosome-wide level was mapped to SSC18 in the same region as the SSC18 OR QTL, were no other TBA QTL have been reported previously. Tribout *et al.* (2008) detected a QTL at the other end of this chromosome in a LW x French LR population. The number of QTL for this trait mapped in other studies is low, and there is little agreement between studies (Cassady *et al.*, 2001; Tribout *et al.*, 2008; Li *et al.*, 2009). However, several associations of this trait with different genes across a range of chromosomes have been explored previously (Rothschild *et al.*, 1996; Buske *et al.*, 2005; Horogh *et al.*, 2005; Buske *et al.*, 2006b; Muñoz *et al.*, 2010; Fernandez-Rodriguez *et al.*, 2010), none located in the QTL found in the present study. The associations between TBA and polymorphisms in the ESR locus on SSC1 reported by Rothschild *et al.* (1996) and validated with much larger datasets by the same research group could not be confirmed in The Roslin populations (Gibson *et al.*, 2002). The absence of any TBA QTL in the current analyses confirms the earlier report from Gibson *et al.* (2002).

Similar results were found for LS, with a low number of mapped QTL. The overlap between these two traits is important and appreciable, as apparent from comparisons

of QTL mapped for these traits. Both, QTL and gene association positions, overlaps greatly between LS and TBA (Rothschild *et al.*, 1996; Buske *et al.*, 2005; Horogh *et al.*, 2005; Li *et al.*, 2009; Fernandez-Rodriguez *et al.*, 2010).

In the present study, a QTL with effects on LS significant at the 5% chromosome-wide level was mapped to SSC6. In a previous study, Wilkie *et al.* (1999) mapped a QTL to the same region of this chromosome for this trait in a Yorkshire x MS population. On this same chromosome, Yasue *et al.* (1999) identified 20 genes in a region of 7 cM associated with LS, and these genes include the pregnancy-specific beta 1-glycoprotein gene. Recently, another candidate gene (*SULT2A1*) for this trait was mapped by Fernandez-Rodriguez *et al.* (2011). However, this gene was not in the region for the QTL mapped in the present study, but in the region of the QTL mapped by Noguera *et al.* (2009) in a bi-dimensional analysis for this trait, where they studied the epistasis between QTL on different chromosomes.

For LS, a QTL on SSC18 was detected in the same region as the QTL for OR and TBA in this study. The number of QTL with effects on reproductive traits which have been mapped to this chromosome is small, but there are some credible candidate genes, including *INHBA* and *IGFBP1*. In a recent study, Sironen *et al.* (2010) mapped the *IGFBP1* gene to SSC18 and tested *IGFBP1* polymorphisms for associations with reproduction trait in a Finnish Yorkshire and Landrace population. A positive effect of one allele of one SNP on LS in later parities of Landrace sows was found (Sironen *et al.*, 2010). The IGFBP1 protein is involved in regulating the menstrual cycle, ovulation, implantation, and foetal growth in humans (Fowler *et al.*, 2000). This gene is also known to play an important role in prenatal development and cell movement (Wang *et al.*, 2006). In a recent study, Miese-Looy *et al.* (2012) inspected the expression of IGFBP1 protein by immunohistochemistry in reproductive tissue, finding the presence of this protein in the endometrium. All these factors and the mapping of this gene in the area of the QTL found in the present study for TBA, LS and OR make *IGFBP1* a good candidate gene for MAS.

There are other traits related to LS which were not analysed in the present study that should be taken into consideration in the analysis of LS as a composite trait. For

NFF, a single QTL mapping to SSC11 was observed in two studies performed in the same population (Cassady *et al.*, 2001; Holl *et al.*, 2004). NVE, as observed in gilts slaughtered at 30 days of gestation, which could be predictive of LS or TBA, was recorded by Bidanel *et al.* (2008) in a study with a LW x MS population. In this study, QTL close to the ones found in the present study were mapped for other traits. As a result they found QTL for NVE on SSC6, SSC9, SSC12 and SSC18. Furthermore, the QTL for NVE on SSC6 which was significant at the 5% chromosome-wide level was mapped close to the QTL for LS found in this study. Similar results were found when comparing the position of the QTL on SSC18 with the one(s) in the present study for TBA, LS and OR. Bidanel *et al.* (2008) suggested Leptin gene (*LEP*) as a candidate gene in this chromosome, but it was discarded as a candidate gene as maps far from the QTL. Number of stillborn (NSB) and number of mummified piglets (NMUM) are both important traits related with the PS trait which was analysed in the present study but no QTL apart from the one in SSC8 were mapped. The only study in which QTL with effects on NMUM have been mapped was on a LW x LR population selected for OR and ES by Holl *et al.* (2004), and QTL on SSC2, SSC6, and SSC12 were found. For NSB, a total of 18 QTL have been mapped on SSC4, SSC5, SSC6, SSC7, SSC8, SSC11, SSC12, SSC13, SSC14, and SSC17 in different studies (Wilkie *et al.*, 1999; Cassady *et al.*, 2001; Holl *et al.*, 2004; Andersson & Georges, 2004; Tribout *et al.*, 2008; Li *et al.*, 2009; Stinckens *et al.*, 2010).

Some of the QTL detected in the present study have not been reported previously. The diversity of results between the different studies illustrates the genetic variation in the different populations. This conclusion is consistent with the greater degree of agreement between studies in which similar populations are used. One of the important factors determining the power of QTL studies is the number of animals that form the population. Increasing the number of traits recorded and the number of genotyped animal is the most effective way of improving the confidence in the findings. Although the number of genotyping assays available has increased with the advent of SNP chips (Ramos *et al.*, 2009) and the cost of genotyping has reduced dramatically, the cost of acquiring phenotypes remains a challenge, especially for

traits such as OR and PS. Thus, it remains difficult to identify genes to improve reproductive traits with equal effects on the different existing breeds, especially for composite traits like LS (Bennett & Leymaster, 1989), expressed by the embryo and the dam (Linville *et al.*, 2001) and influenced by environmental factors.

Chapter 3
Analysis of chromosome 8

3.1. Introduction

In the previous chapter, a scan of the pig genome for QTL with effects on female reproductive performance traits is described. Although the putative QTL identified on SSC8 are listed in the tabulated results, these findings are not discussed. The analysis of SSC8 is discussed here in greater depth, as The Roslin LW-MS populations have been examined earlier for QTL on SSC8 with effects on reproductive traits (King *et al.*, 2001; King *et al.*, 2003; King, 2003). The earlier analysis found evidence for the presence of a QTL on the q arm of SSC8 controlling LS and PS in animals at first parity. The effects of both these QTL were negative overdominant, with no significant additive effects. A putative QTL for number of teats (TN) was also found. The QTL for PS and TN were significant at the 5% chromosome-wide level and the LS one was significant at the 1% chromosome-wide level. The QTL for LS and PS were co-located at the end of the q arm of the chromosome.

One of the original reasons to search for reproductive QTL on SSC8 was that the pig homologue of the sheep Booroola fecundity gene (FecB), which has been shown to be *BMPR1B* (Wilson *et al.*, 2001), was predicted to map on the q arm of SSC8. The Booroola gene is known to improve LS in sheep through increases in OR. However, there is no comparable evidence that *BMPR1B* is a major gene for prolificacy in pigs (Kim *et al.*, 2003; Tomas *et al.*, 2006; Casellas *et al.*, 2008), and there is no evidence for a QTL for OR in the region of the pig *BMPR1B* gene.

The gene encoding secreted phosphoprotein 1, also known as osteopontin (*SPP1*), lies within the 95% confidence interval of the LS and PS QTL reported by King *et al.* (2003). As *SPP1* also has an important role in embryo implantation and placentation (Nomura *et al.*, 1988; Johnson *et al.*, 1999a; Johnson *et al.*, 1999b), it is not only a positional, but also a physiological candidate gene for reproductive traits. Furthermore, the presence of a SINE in the *SPP1* gene has been reported to be associated with LS in second and subsequent parities (Knoll *et al.*, 1999).

Although there is evidence from several independent studies suggesting *SPP1* as a physiological and positional candidate gene involved in the variation seen in PS and

LS among pig breeds, the confidence interval for the LS and PS QTL on SSC8qter is large. Therefore, another gene or genes within the QTL region identified on SSC8qter may be responsible for the observed variation in LS and PS. The presence of multiple QTL in this region could also indicate the presence of another gene close to *SPP1* with function in reproduction. The emerging pig genome sequence (Archibald *et al.*, 2010) provides a valuable source of information on the gene content of the QTL, but the annotation of the genome sequence was incomplete at the time this study was performed.

3.2. Materials and Methods

3.2.1. Population, Phenotypic trait data and DNA samples

The population used for this study was the same as the one used for the whole genome analysis. The population structure, phenotypic trait data and DNA samples preparation are described in Chapter 2 (2.2.1, 2.2.2, 2.2.3).

3.2.2. Genotyping of microsatellite markers

The genotypes of the F_2 trait-recorded females, their F_1 parents, and their purebred grandparents were determined for 13 additional polymorphic genetic markers located mainly in the region of the QTL discovered previously in this population, in order to increase marker density. Initially, ten markers were genotyped, which included six selected from the USDA-MARC linkage map (KS192, KS904, SW763, SW790, SW1551, SW1980) (http://www.marc.usda.gov/genome/swine/swine.html) and four microsatellites (S0782, S0792, S0793, S0794) developed from BAC end sequences of BAC clones (PigE-139L4, PigE-55F17, PigE-115B2, PigE-190O14, respectively) that map to the region of interest in the physical map (Humphray *et al.*, 2007); (http://pre.ensembl.org/Sus_scrofa_map/Info/Index). In a second round of fine mapping, three microsatellite markers designed from BAC clone sequences (CH242-443f10, CU467102; CH242-238o22, CU606871; CH242-27o17, CU633175) were genotyped. The information for these markers is presented in Table 3.1. The genotyping was performed as described in Chapter 2 (2.2.4). For the first set of

primers TAQ (Roche) was used and for the second set FastStart TAQ (Roche) was used.

Marker name	Primers	Primer sequence 5' - 3'	Product size bp	Tm
S0782	bT139L4SP6-FAM	GAGGGTGAGAGAGTCAGAGGAGA	117-167	57
	bT139L4SP6-R	GATGGTTTCCTGGAAGCAGAGCTA		
S0792	bT55F17T7-VIC	GGAATATTTCACGAG ATGCT CAA	156-192	62
	bT55F17T7-R	CCCTAGCCTGAGAACCTCCACAT		
S0793	bT115B2T7-PET	CAAGGTGGGAACCAGGCATACATA	111-146	55
	bT115B2T7-R	TCTGTAATCATTTACTGTGGGTGACCA		
S0794	bT190O14T7-NED	TTCTGCTGCTCAATATTGGACGTT	236 -243	57
	bT190O14T7-R	TGGCTGATTCTTGTGAACTGTGA		
KS192	KS192-NED	GGAGACTTGTGGTTTAACTGGC	165-199	60
	KS192-R	TCCTACTGCTTCCCAACCC		
KS904	KS904-PET	AAAACCCTGGGCTGAGATG	160-173	55
	KS904-R	CAATGAAATGGGGAAGAAGC		
SW763	SW763-FAM	GGGTGCATTGTTCTCATATGG	156-189	62
	SW763-R	TGCTCTAGCAACACACACCC		
SW790	SW790-VIC	CTGTGGGAGTGTAGCATCTTTG	111-186	58
	SW790-R	CATACACCCCAGATGTGGC		
SW1551	SW1551-PET	TTTACTTGGGGAAACCCTCC	150-188	55
	SW1551-R	GATCAACCCAAATTCTTGGC		
SW1980	SW1980-NED	GCTTCTGTATGCCACAGCTG	174-200	58
	SW1980-R	CCCCCATTTGAACAATGAAG		
238o22	238o22-PET	CCAAGGCCGTGTGTGAGGATTAT	199	55
	238o22-R	TGGAAAAACACTTCAGGCAACTG		
27o17	27o17-NED	CCTTCTCTCTCCCATTTCCTTCTC	195	58
	27o17-R	TCCTTCTCTCCTTCTTTGCCTTTC		
443f10	443f10-PET	TCTTTCCAAGGGATCATAAAGTCTGA	143	50
	443f10-R	GGCTCTCTGATCCCAAATCCTGA		

Table 3.1 List of markers. This Table includes name of the markers, fluorescent label for each forward primer, sequence of forward and reverse primers, size of the product amplified by the primers in bps and optimal annealing temperature (Tm) for each pair of primers.

3.2.3. Linkage map construction

The initial linkage map (Map 2) was built as described in Chapter 2 (2.2.5), using the ten markers genotyped together with other markers genotyped previously in this

population (Appendix 1). This included microsatellite and gene-associated markers (Table 3.2) used in the previous study of this chromosome (Map 1), and genotyped at The Roslin Institute or by collaborators at the University of Guelph, Canada. A haplotype of genetic markers in the *KIT* locus, which was developed for an earlier study of this locus and its effects on coat colour, was added to the list of markers to include in the map. The information for these markers was exported from ResSpecies. A second map (Map 3) was built with the markers used for the first map and the three microsatellites genotyped in the second stage of this study.

Marker	Marker type	Reference
Anonymous DNA markers		
S0017	Microsatellite	(Coppieters *et al.*, 1993)
S0178	Microsatellite	(Ellegren *et al.*, 1994)
S0225	Microsatellite	(Robic *et al.*, 1994)
SW7	Microsatellite	(Rohrer *et al.*, 1994)
SW61	Microsatellite	(Rohrer *et al.*, 1994)
SW268	Microsatellite	(Rohrer *et al.*, 1994)
SW905	Microsatellite	(Rohrer *et al.*, 1994)
SW2410	Microsatellite	(Alexander *et al.*, 1996)
SW2611	Microsatellite	(Alexander *et al.*, 1996)
Gene-associated markers		
AREG	PCR-RFLP	(Jiang *et al.*, 2002b)
FGG-2	Bi-PASA	(Jiang *et al.*, 2002b)
IBSP-1	Bi-PASA	(Jiang *et al.*, 2002a)
GNRHR-1	Bi-PASA	(Jiang *et al.*, 2001)
GNRHR-2	Bi-PASA	(Jiang *et al.*, 2001)
HD-1 (HTT)	PCR-RFLP	(Jiang *et al.*, 2002b)
QDPR1	PCR-RFLP	(Jiang *et al.*, 2002b)
SLIT2-1	PCR-RFLP	(Jiang *et al.*, 2002b)
SPP1-1	Bi-PASA	(Moran, 1993)
SPP1-4	PCR-RFLP	(Knoll *et al.*, 1999)
SPP1-5	Bi-PASA	(Jiang *et al.*, 2002a)
STE-1 (SULTE-1)	PCR-DSCP	(Jiang *et al.*, 2002b)

Table 3.2 Information for markers for SSC8 from previous study. The Table shows marker name, marker type, and the reference in which details of the primers have been reported.

3.2.4. QTL scan analysis

The scan of this chromosome for QTL with effects on reproductive traits was performed as described in Chapter 2 (2.2.6) for all the traits, i.e., TBA, LS, PS, OR and TN, with the difference that a chromosome-wide analysis was run instead of a genome-wide analysis. The analysis was first performed for the map built in the previous study of this chromosome, indicated from now on as Map 1 (King *et al.*, 2003), using the same phenotypic files prepared for the present study, for Age Group 1 (AG1). Then, the analysis for the first map of this chromosome developed in the current study (i.e., Map 2) was performed for AG1 and AG2 animals. For the latter group, parity (1 or 2) was fitted as a fixed effect in the chromosome-wide analysis. The second map from the present study (i.e., Map 3) was explored only for AG1 animals. A bootstrap with resampling was also run for all the analyses in order to define the confidence intervals for the QTL.

For this chromosome, when evidence for a QTL was found, the presence of a second QTL in the same chromosome was investigated. The best 'two-linked-QTL' model was identified by a grid search at 1 cM resolution of all possible positions for two QTL. The best-fitting model with two QTL was tested against the best model fitting only one QTL using an F-test. The *F*-ratio was calculated by $[(RSS1 - RSS2)/(df1 - df2)]/(RSS2/df2)$ with $(df1 - df2)$ degrees of freedom in the numerator considering additive and dominance effects in the genetic model. The two-QTL model is accepted if there is a significant improvement over the best one-QTL model at $P < 0.05$.

3.2.5. Comparative maps

Using the standard nucleotide-nucleotide BLAST search, the human genome sequence was searched for sequence homology to the sequences of the markers used to build the linkage map (Map 2), and a comparative map between SSC8 and HSA4 was established. The comparative map, matching markers by names, was drawn with the ArkDB mapping option (ArkMAP, The Roslin Institute; http://www.thearkdb.org/arkdb/download.jsp) with a tool where the maps with the marker positions were introduced, for SSC8 in cM and for HSA4 in Mbps, and the

name of the markers, which represent homologous sequences, were linked with a line.

For the second map (Map 3), the position of each marker in the linkage map in cM was compared to the position of these markers in Mbp in the *Sus scrofa* sequence for SSC8 (Sscrofa9; http://www.ensembl.org/Sus_scrofa/Info/Index). First the draft pig genome sequence (Sscrofa9) was searched for matches to the sequence of each marker. Once the match was found, the Mbp position was annotated next to the marker. The comparative map was drawn using the ArkMAP, a desktop Map drawing Tool (Java Web Start), where the maps were compared, and the markers linked by name.

3.3. Results

3.3.1. Phenotypic data

The phenotypic data used in the SSC8 analyses are the same as those used for the genome scan, described in Chapter 2 (2.3.1).

3.3.2. 1st genotyping

3.3.2.1. Microsatellite genotyping and Linkage map

The genotyping results from the ABI 3730x1 DNA analyser (Applied Biosystems) for the 10 markers were analysed with GeneMapper software (Applied Biosystems) as indicated in Chapter 2 (2.2.4). Results for the markers genotyped in this study are presented in Table 3.3 and information for some of the other markers is presented in Appendix 1.

Marker name	Individuals with results
S0782	299
S0792	307
S0793	302
S0794	298
KS192	303
KS904	289
SW763	300
SW790	307
SW1551	305
SW1980	298

Table 3.3 Number of individuals with genotype results for each marker. The Table shows the list of markers genotyped in this study and the number of individuals with genotype results.

All the marker information was exported from ResSpecies, the **.input** file, to build the map with CRI-MAP and MultiMap, was prepared including the markers to be haplotyped, *GNRHR* and *SPP1*, and the map was built as described in 2.2.5. The resulting linkage map (Map 2) was consistent with the published USDA-MARC linkage map for this chromosome and the resulting map is presented in Table 3.4. The new map (Map 2) consisted of 32 markers covering 148.1 cM compared with the 21 markers covering 139.8 cM on the map from the previous study of this chromosome (Map 1) (King *et al.*, 2003).

	Map 1 (King *et al.*, 2003)	Map 2 this study	USDA map
Marker name	**Position (cM)**	**Position (cM)**	**Relative Position (cM)**
SW2410	0.0	0	-1.3
HD	2.8	2.8	-
SW2611	3.9	3.9	2.5
SW905	21.9	22	20.8
QDPR	38.0	38	-
SLIT2	49.0	49.1	-
SW268	54.3	54.3	27
SW7	71.9	72.3	55.4
KIT	-	78.6	-
GNRHR-1	80.3	80.9	-
GNRHR-2	80.3	80.9	-
SULTE1	80.6	81.2	-
S0017	80.6	81.2	60.4
AREG	81.1	81.8	-
FGG	83.4	84	-
S0225	96.2	97.1	82.8
S0794	-	98.9	-
KS192	-	101.2	89.9
SW763	-	102.3	92.4
S0793	-	107.2	-
SW1551	-	114.1	105.9
SW790	-	116.7	107.5
SW61	115.2	119.7	112.3
S0782	-	123.5	-
SPP1-1	126.8	132.4	120.2
SPP1-4	126.8	132.4	-
SPP1-5	126.8	132.4	-
SPP1-6	-	132.4	-
IBSP	127.0	132.6	-
S0792	-	135.6	-
SW1980	-	137.5	126.1
KS904	-	143.3	125
S0178	139.8	148.1	127.2

Table 3.4 Linkage map for SSC8. The Table shows name of the markers, and position of the markers for the previous map (Map1) followed by the map built in this study (Map 2) and the relative position of the markers in the USDA-MARC map.

3.3.2.2. QTL analysis

The QTL analyses were performed as previously described in Chapter 2 (2.2.6). Map 1 and Map 2 were used in separate analyses of AG1 animals. A chromosome-wide with 1,000 permutation analysis was performed for all traits at the same time, followed by a bootstrap with resampling analysis for each trait with 1,000 permutations. Gestation length, the only covariate with a significant effect on the trait, was used as a covariance for all the traits, except for TN, for which no covariates were used. The results of the analysis for AG1 are presented in Table 3.5 and Figure 3.1 for Map 1 and in Table 3.6 and Figure 3.2 for Map 2. Table 3.7 and Figure 3.3 show results for AG2 with Map 2.

The chromosome-wide analysis for Map 2 for AG1 revealed two significant QTL at the 5% chromosome-wide level for LS and TBA and a significant QTL at the 1% chromosome-wide level for PS. The QTL for TBA was co-located with the LS QTL at 114 cM and the QTL for PS was located at 135 cM. No significant QTL were found for OR and TN.

The analysis performed for the previous map (Map 1) resulted in a significant QTL at 5% chromosome-wide level for PS at 129 cM. For the rest of the traits no significant QTL were found. The analysis for Map 2 did not reveal any significant QTL position for AG2 animals for any of the traits. The background genetic effects analysis for position 114 cM and 135 cM revealed no significant QTL. The results for this analysis are presented in Table 3.8.

SSC8 QTL analysis using Map 1 (King *et al.*, 2003)

Trait	Position (cM)	F-ratio	Estimated Effect		95% CI (cM) (start-end)	Significance threshold	
			Additive effect (±SE)	Dominance effect (±SE)		P <0.05	P<0.01
TBA	115	5.31	0.0047 (0.42)	-2.11 (0.65)	0.0-135.0	5.78	7.80
LS	130	4.13	-0.49 (0.48)	-1.898 (0.72)	2.0-139.0	5.46	7.24
PS	129	7.39	-0.03 (0.02)	-0.118 (0.03)	66.0-137.0	6.03	7.79
OR	0	2.19	0.83 (0.41)	-0.07 (0.58)	0.0-127.0	5.76	7.93
TN	49	2.56	-0.123 (0.17)	0.56 (0.25)	4.5-127.0	5.56	8.07

Table 3.5 Results from chromosome-wide and bootstrap analyses on SSC8 for AG1 animals with Map 1 (King *et al.*, 2003). The Table indicates (by columns) trait analysed, position of the QTL in cMs, F-ratio for the QTL, estimate additive and dominance effect (± Standard error), confidence interval (CI) in cM, and significance threshold for each trait (King *et al.*, 2003). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

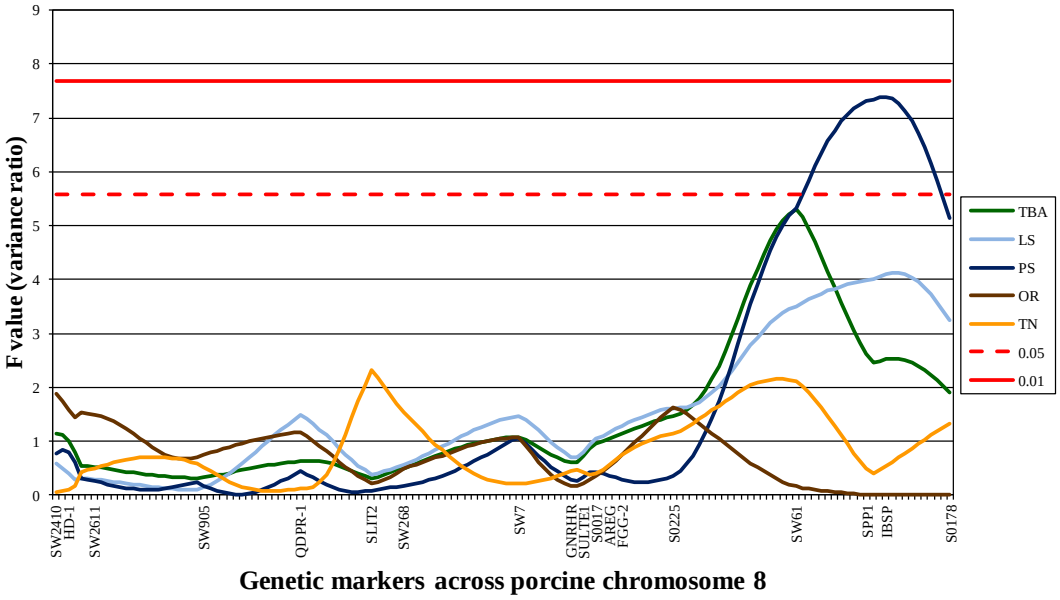

Figure 3.1 Interval mapping of TBA, LS, PS, OR and TN on SSC8 for AG1 animals with Map 1. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

SSC8 QTL analysis using Map 2 – AG1

Trait	Position (cM)	F-ratio	Estimated effect		95% CI (cM) (start-end)	Significance threshold	
			Additive effect (±SE)	Dominance effect (±SE)		P<0.05	P<0.01
TBA	114	7.3	0.04 (0.39)	-2.17 (0.57)	0.5-143.5	5.86	8.33
LS	114	6.12	0.08 (0.44)	-2.25 (0.64)	2.0-146.5	5.81	8.14
PS	135	7.48	-0.03 (0.02)	-0.108 (0.03)	77.0-172.5	5.74	7.38
OR	0	2.8	0.76 (0.38)	-0.74 (0.57)	0.0-142.5	5.89	7.66
TN	123	3.23	0.22 (0.15)	-0.48 (0.22)	4.0-147.0	6.16	8.22

Table 3.6 Results from chromosome-wide and bootstrap analyses on SSC8 for AG1 animals with Map 2. The Table indicates trait analysed, position of the QTL in cMs, F-ratio for the QTL, estimate additive and dominance effect (± Standard error), confidence interval (CI) in cM, and significance threshold for each trait. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

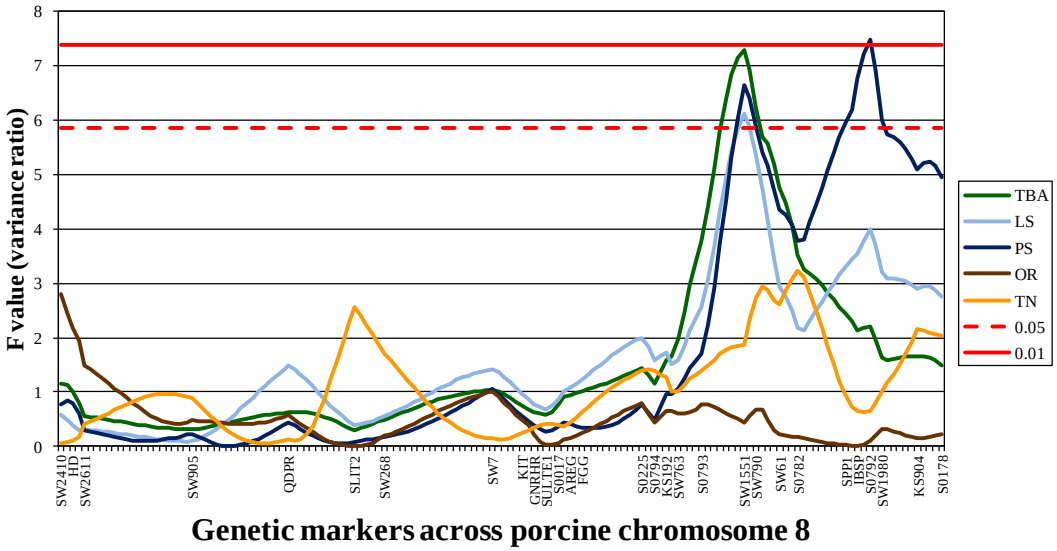

Figure 3.2 Interval mapping of TBA, LS, PS, OR and TN on SSC8 for AG1 animals with Map 2. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

SSC8 QTL analysis using Map 2 – AG2

Trait	Position cM	F-ratio	Estimated effect		95% CI (cM) (start-end)	Significance threshold	
			Additive effect (±SE)	Dominance effect (±SE)		P<0.05	P<0.01
TBA	38	2.01	-0.399 (0.49)	-1.11 (0.66)	0.0-147.0	5.99	7.91
LS	123	1.44	-0.11 (0.44)	1.072 (0.64)	0.0-145.0	5.88	7.12
PS	38	2.76	-0.03 (0.02)	-0.064 (0.03)	0.0-143.0	5.80	7.41
OR	147	1.23	0.18 (0.51)	-1.19 (0.8)	0.5-147.0	5.84	7.99
TN	97	1.78	-0.04 (0.17)	0.45 (0.24)	4.0-147.0	5.69	7.79

Table 3.7 Results from chromosome-wide and bootstrap analyses on SSC8 for animals AG2 with Map 2. The Table indicates trait analysed, position of the QTL in cMs, F-ratio for the QTL, estimate additive and dominance effect (± Standard error), confidence interval (CI) in cM, and significance threshold for each trait. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

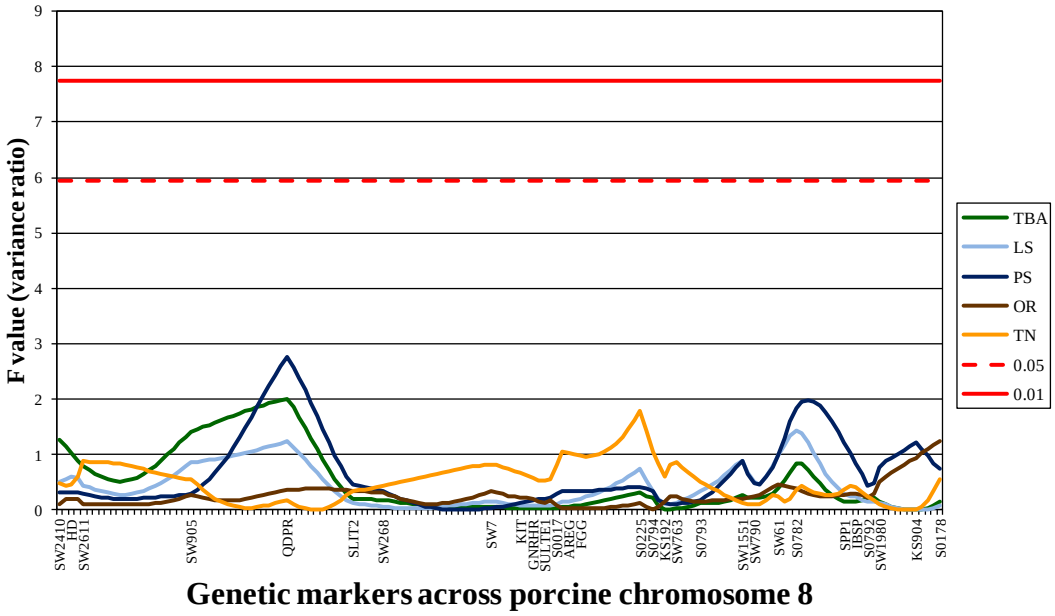

Figure 3.3 Interval mapping of TBA, LS, PS, OR and TN on SSC8 for AG2 animals with Map 2. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

Position fitted	Trait	Position cM	*F*-ratio	Estimate effect		Significance threshold	
				Additive effect (±SE)	Dominance effect (±SE)	P<0.05	P< 0.01
114 cM	**TBA**	0	1.39	0.26 (0.39)	-0.86 (0.57)	6.04	8.68
	LS	136	3.36	-1.49 (0.61)	-0.66 (0.70)	5.61	8.14
	PS	135	4.07	-0.05 (0.03)	-0.07 (0.03)	6.05	5.58
135 cM	**TBA**	114	5.56	0.50 (0.57)	-2.07 (0.64)	5.93	7.66
	LS	114	5.37	1.18 (0.64)	-1.93 (0.71)	6.20	8.20
	PS	114	3.27	0.03 (0.03)	-0.08 (0.03)	5.65	7.52

Table 3.8 Results of the chromosome-wide analysis when QTL at position 114 cM and 135 cM were fitted as background genetic effects. TBA, total number born alive; LS, litter size; PS, prenatal survival.

3.3.2.3. Comparative map to Human

The comparative map between the linkage map (cM) for SSC8 and the human sequence map (Mbp) for HSA4, is presented in Figure 3.4. This comparative map confirms the homology between SSC8 and HSA4. The order and orientation of the regions of homology clearly differ between HSA4 and SSC8, as shown by the major chromosomal inversion between FGG and S0178 on SSC8 relative to the orientation of this region on HSA4. The pattern of homology between SSC8 and the human genome is more complex than this comparative map reveals (see Meyers *et al.*, 2005; Vingborg *et al.*, 2009 and comparative 'synteny views' in the Ensembl genome browser). However, for the region of interest towards the end of the long arm of SSC8, the gene content of HSA4 between 82 and 156 Mbp is expected to be similar to the gene content of SSC8 between FGG and S0178 but with gene order inverted relative to one another.

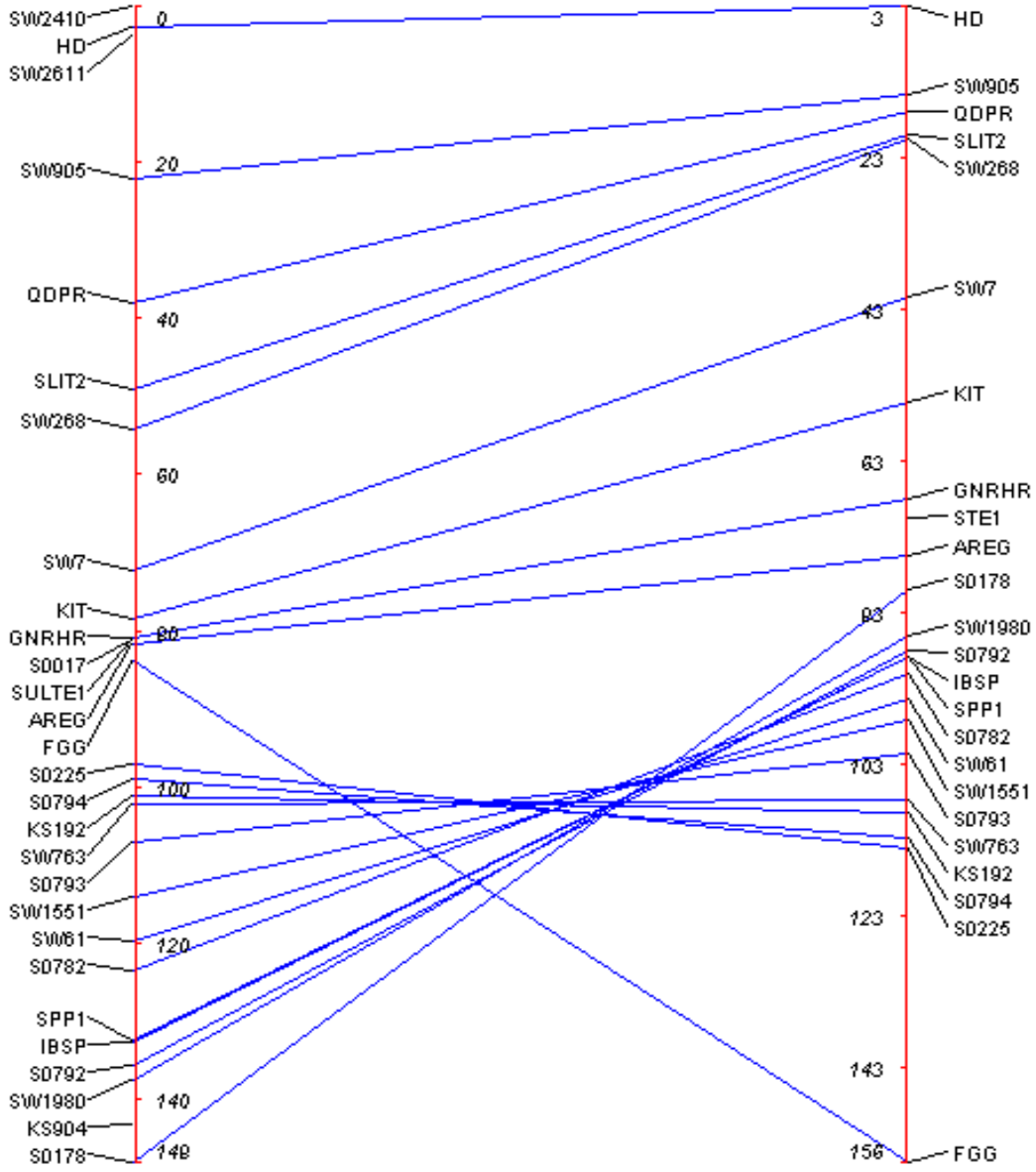

Figure 3.4 Comparative map of SSC8 (left) linkage map (Map 2) in cM and the sequence of HSA4 in Mbp (right).

3.3.3. 2nd genotyping

3.3.3.1. Genotyping and Linkage map

The population was genotyped for three additional markers as previously described in Chapter 2 (2.2.4) using the FastStart Taq protocol. The number of animals with results for these markers, presented in Table 3.9, was between 299 and 304.

Marker name	Individuals with results
CH242-238o22	304
CH242-27o17	299
CH242-443f10	304

Table 3.9 Number of individuals with genotype results for each marker. The Table shows the list of markers genotyped in this study and the number of individuals with genotype results.

These three markers together with the previous ones were used to build a second map, as described in 2.2.5. The result for this map (Map 3) is shown in Table 3.10. Map 3 has a total of 36 markers covering 136.0 cM. The 238o22 and 27o17 markers provided additional resolution in the S0793-SW1551 interval which was 6.8 cM in Map 2. The 443f10 marker provided some additional resolution in the SW1980-KS904 interval.

3.3.3.2. QTL analysis

The chromosome-wide and bootstrapping QTL analyses of SSC8 were repeated for the AG1 animals using linkage Map 3. The results are presented in Table 3.11 and Figure 3.5. Also results of the permutation and bootstrapping analyses are presented in Figures 3.5 to 3.8 for the three traits where significant QTL were found, and just for the region of the QTL.

Marker Name	Position (cM)	Kosambi distance (cM)
SW2410	0.0	0.0
HD-1	2.5	2.5
SW2611	3.7	1.2
SW905	19.3	15.6
QDPR	32.6	13.3
SLIT2	44.1	11.5
SW7	65.2	21.1
KIT	70.7	5.5
GNRHR-1	72.8	2.1
GNRHR-2	72.8	0.0
SULTE1	73.1	0.3
S0017	73.1	0.0
AREG	73.7	0.6
FGG-1	75.6	1.9
FGG-2	75.6	0.0
S0225	89.1	13.5
S0794	90.9	1.8
KS192	93.5	2.6
SW763	94.1	0.6
S0793	98.7	4.6
238o22	101.3	2.6
27o17	103.0	1.7
SW1551	105.0	2.0
SW790	106.1	1.1
SW61	108.8	2.7
S0782	112.3	3.5
SPP1-1	121.2	8.9
SPP1-4	121.2	0.0
SPP1-5	121.2	0.0
SPP1-6	121.2	0.0
IBSP	121.4	0.2
S0792	124.4	3.0
SW1980	125.9	1.5
443f10	130.9	5.0
KS904	131.3	0.4
S0178	136.2	4.9

Table 3.10 Linkage map for SSC8 with three more markers (highlighted). The Table includes name of each marker, position for markers in cM, and the Kosambi distance between markers.

SSC8 QTL analysis using Map 3 – AG1

Trait	Position cM	F-ratio	Estimate effect		95% CI (cM) (start-end)	Significance levels	
			Additive effect (±SE)	Dominance effect (±SE)		P<0.05	P<0.01
TBA	105	6.98	-0.03 (.038)	-2.12 (0.56)	0.0-133.0	5.60	7.36
LS	105	5.86	-0.39 (0.43)	-2.18 (0.63)	1.5-135.0	5.62	7.88
PS	124	7.53	-0.03 (0.02)	-0.1 (0.03)	2.0-136.0	5.83	7.53
OR	0	2.23	0.84 (0.4)	-0.09 (0.57)	0.0-136.0	5.79	8.26
TN	112	3.31	0.21 (0.15)	-0.49 (0.23)	5.5-131.0	6.00	7.47

Table 3.11 Results from chromosome-wide and bootstrap analyses on SSC8 for AG1 animals with Map 3. The Table indicates trait analysed, position of the QTL in cMs, F-ratio for the QTL, estimate additive and dominance effect (± Standard error), confidence interval (CI) in cM, and significance threshold for each trait. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

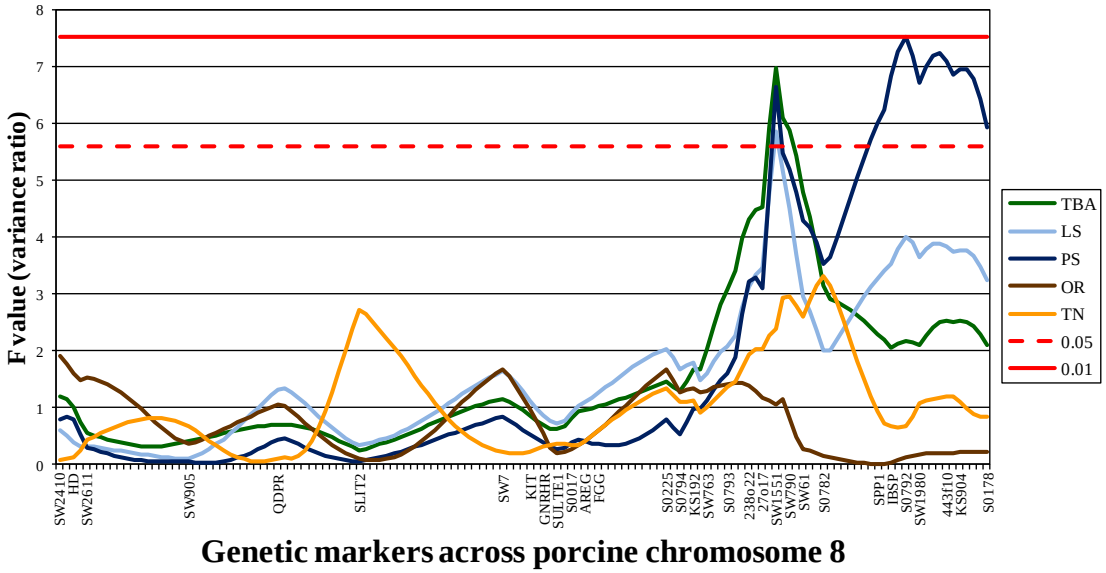

Figure 3.5 Interval mapping of TBA, LS, PS, OR and TN on SSC8 for AG1 animals with Map 3. Chromosome-wide significance level at P<0.05 (broken red line) and P<0.01 (solid red line). TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

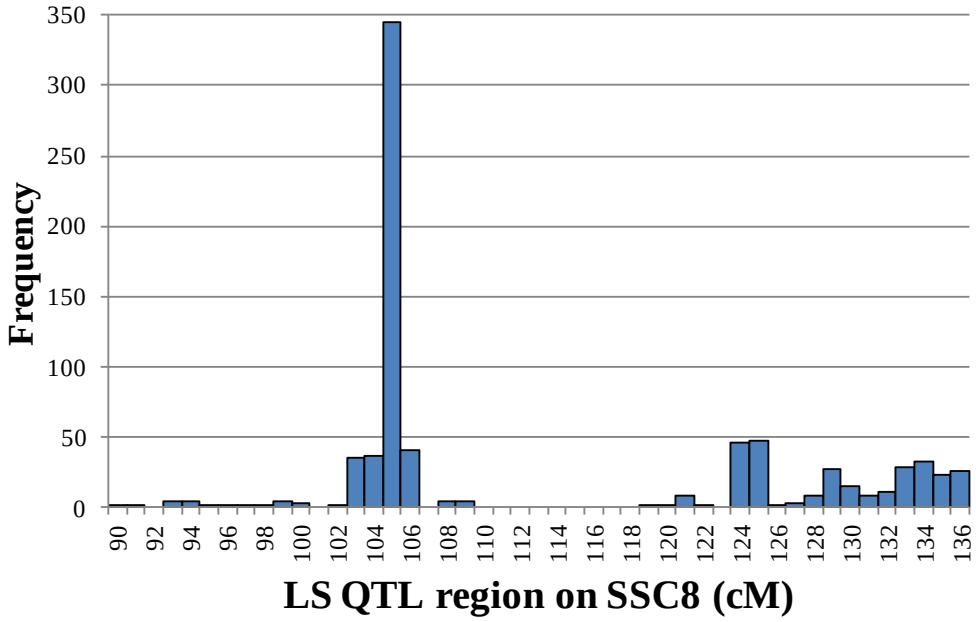

Figure 3.6 Permutation and bootstrapping analyses results for the litter size analysis for Map 3. The Figure shows results only for the region where the QTL was mapped.

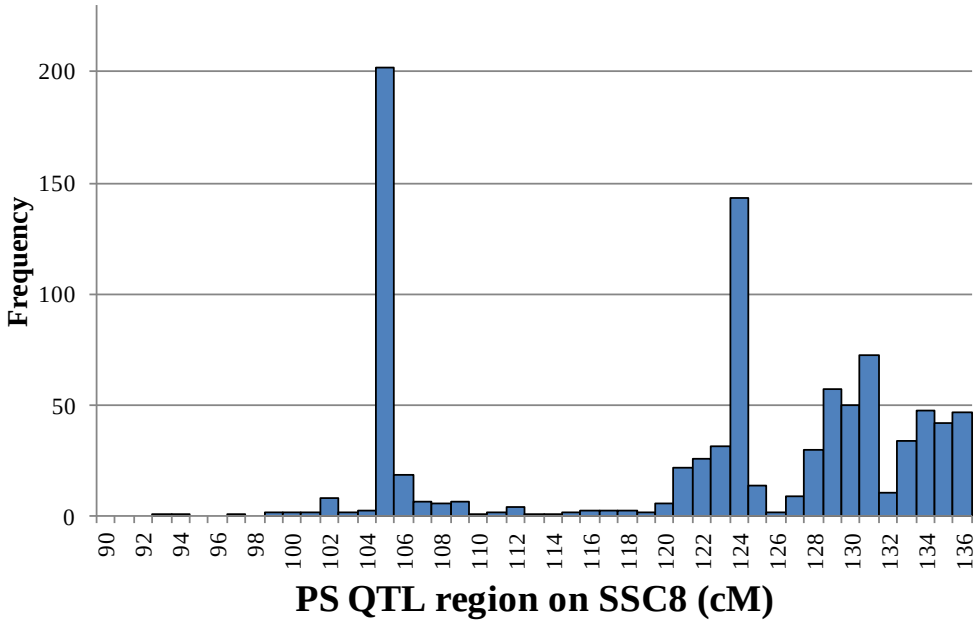

Figure 3.7 Permutation and bootstrapping analyses results for the prenatal survival analysis for Map 3. The Figure shows results only for the region where the QTL was mapped.

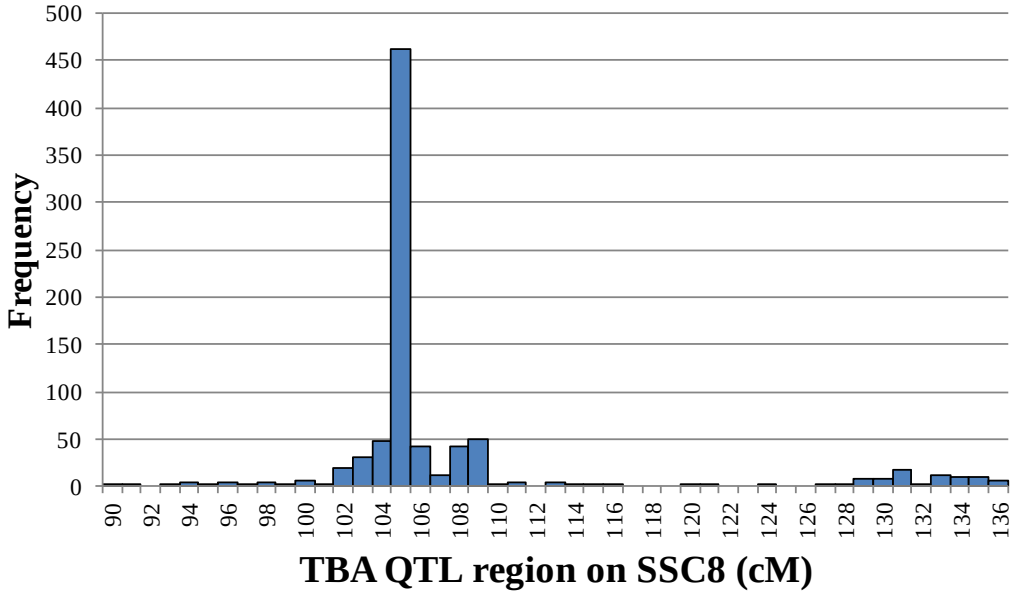

Figure 3.8 Permutation and bootstrapping analyses results for the total number born alive analysis for Map 3. The Figure shows results only for the region where the QTL was mapped.

The chromosome-wide analysis revealed two putative QTL, which were significant at the 5% chromosome-wide level for LS and TBA, and a putative QTL, which was significant at the 1% chromosome-wide level for PS. As in the first analysis, the QTL for TBA was co-located with the LS QTL at 105 cM and the QTL for PS was located at 124 cM. No significant QTL were found for OR and TN. When the QTL at positions 105 cM and 124 cM were included as background genetic effects in the analyses, no further significant QTL were observed. The results for this analysis are presented in Table 3.12 and Table 3.13.

Trait	Position cM	F-ratio	Estimated effect		Significance levels	
			Additive effect (±SE)	Dominance effect (±SE)	P<0.05	P<0.01
TBA	0	2.0	0.354 (0.39)	-1.022 (0.56)	5.73	7.37
LS	124	3.04	-1.41 (0.63)	-0.811 (0.71)	5.68	7.87
PS	124	3.38	-0.047 (0.03)	-0.040 (0.03)	5.76	7.4
OR	65	1.96	-0.316 (0.49)	-1.183 (0.65)	5.56	7.45
TN	112	3.31	0.21 (0.15)	-0.49 (0.23)	5.96	7.75

Table 3.12 Results of chromosome-wide analysis for all traits with QTL at 105 cM fitted as background genetic effect. The Table shows name of the trait analysed, position of the QTL in cMs, F-value, estimated additive and dominance effect (± Standard error), and significance threshold for each trait. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

Trait	Position cM	F-ratio	Estimated effect		Significance levels	
			Additive effect (±SE)	Dominance effect (±SE)	P<0.05	P<0.01
TBA	105	5.26	0.425(0.57)	-1.992 (0.63)	5.68	8.61
LS	105	4.84	1.033(0.63)	-1.819(0.70)	6.01	7.73
PS	105	2.56	0.014(0.03)	-0.075(0.03)	5.94	8.10
OR	101	3.34	1.430(0.56)	-0.282(0.63)	5.81	8.12
TN	125	4.83	2.619(1.72)	4.724(1.91)	5.69	7.99

Table 3.13 Results of chromosome-wide analysis for all traits with QTL at 124 cM fitted as background genetic effect. The Table shows name of the trait analysed, position of the QTL in cMs, F-value, estimated additive and dominance effect (± standard error), and significance threshold for each trait. TBA, total number born alive; LS, litter size; PS, prenatal survival; OR, ovulation rate; TN, teat number.

The data were examined for evidence of a second QTL for PS and LS. The results for the analyses of 2 QTL *versus* 1 QTL gave an F-value of 2.56 for PS and of 3.04 for LS (Table 3.14 and Table 3.15). The F-distribution Table was inspected for the significance level values at P<0.05 and no evidence was found for an improvement of the 2 QTL model over the 1 QTL model for any of the traits investigated.

Trait	QTL 1 (cM)	QTL 2 (cM)	F 4*df* (2 QTL vs. 0 QTL)	F 2*df* (2 QTL vs. 1 QTL)
LS	105	124	4.54	3.04

Table 3.14 Results of the analysis for a second QTL on SSC8 for litter size (LS) for positions 105 cM and 124 cM. F-value with 4 degrees of freedom (*df*) for the model of 2 QTL *vs.* 0 QTL and for 2 for the model of 2 QTL *vs.* 1 QTL.

Trait	QTL 1 (cM)	QTL 2 (cM)	F 4*df* (2 QTL vs. 0 QTL)	F 2*df* (2 QTL vs. 1 QTL)
PS	105	124	5.14	2.56

Table 3.15 Results of the analysis for a second QTL on SSC8 for prenatal survival (PS) for positions 105 cM and 124 cM. F-value for 4 degrees of freedom (*df*) for the model of 2 QTL *vs.* 0 QTL and for 2 for the model of 2 QTL *vs.* 1 QTL.

3.3.3.3. Comparative map with pig genome sequence

At the time of this second analysis, a draft of the pig genome sequence (Sscrofa9) was available and the position of the markers in cM in the linkage map (Map 3) was compared with the position of these markers in the pig genome sequence. The results are shown in Figure 3.9.

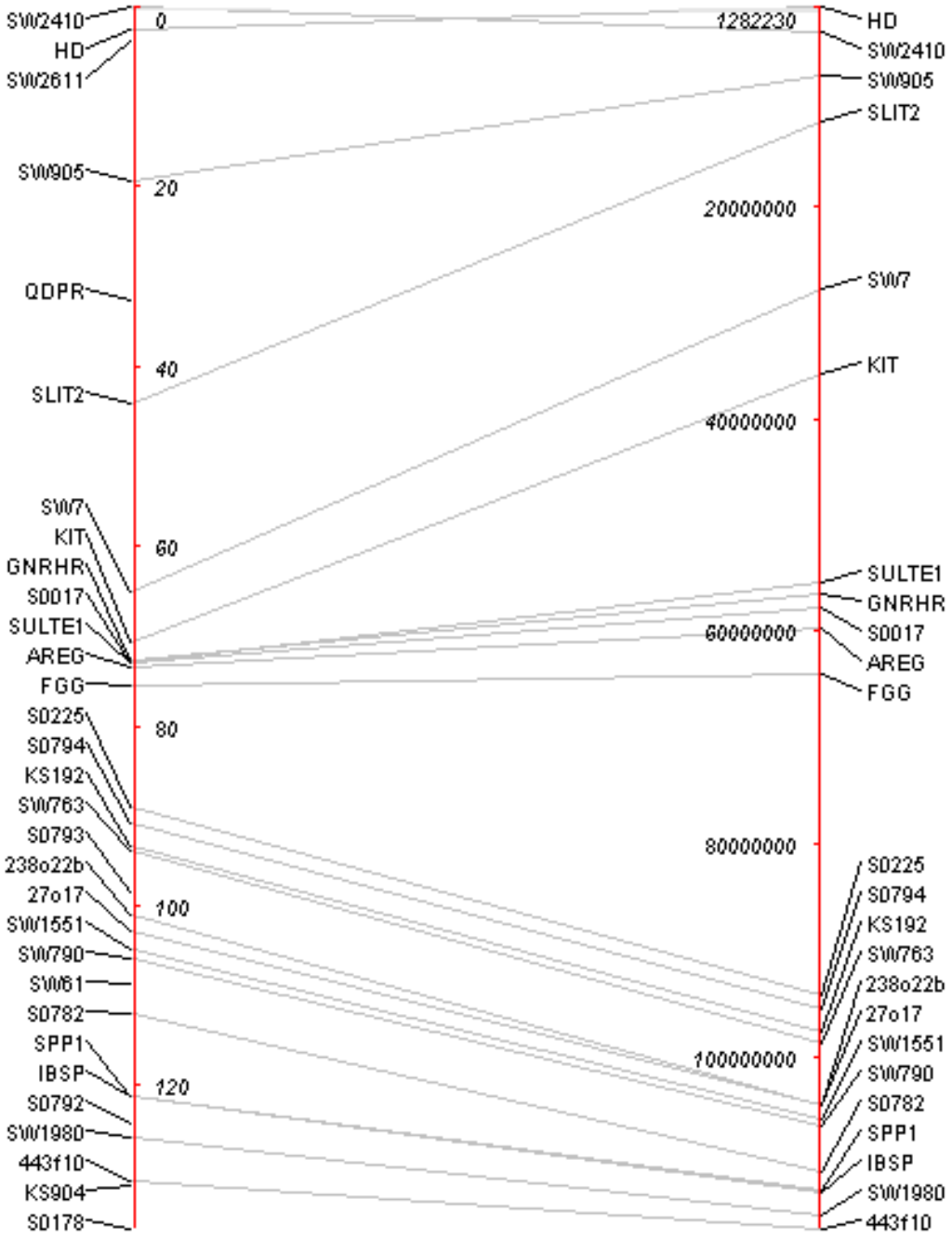

Figure 3.9 Comparative map of linkage map (Map 3) in cM (left) and the sequence map in Mbps (right) for SSC8.

3.4. Discussion

A more detailed scan of SSC8 for QTL with effects on female reproductive traits was conducted in this study, as an earlier study of the same population used here had revealed putative QTL with effects on PS, LS and limited evidence for a QTL with effects on TN (King *et al.*, 2003). For the current study, the population was genotyped for 13 additional genetic markers in the QTL region of interest, in order to improve the resolution with which the QTL were mapped. Chromosome 8 was scanned for QTL with effects on each of the five traits, one of which (TBA) was not analysed in the earlier study. In the current study, three putative QTL which are significant at the 5% (two of them) and at the 1% (one of them) chromosome-wide level were found for age group 1 (animals of parity 1), two of which (LS and PS) were found in the previous study. Age group 2, with records from parity 1 and parity 2, was also analysed for these traits on SSC8 but not significant QTL were found. Different parities have been considered as repeat records of the same trait. However, results from a number of studies demonstrate that different genes control the traits in the different parities (Hanenberg *et al.*, 2001; Noguera *et al.*, 2002; Serenius *et al.*, 2003; Fernandez *et al.*, 2008). Genes with effect only in first parity have a large effect on the value of the QTL and any marker derived from it.

Although the peak positions for the PS and LS QTL reported earlier were 2 cM apart, within the limited accuracy of the QTL mapping in this population, these two QTL were co-located (see Figure 3.5 and Fig 2 in King *et al.*, 2003). The QTL analyses conducted in the present study with Map 3 reveals evidence for QTL with effects on TBA and LS, with a peak position at 105 cM (Figure 3.5, Table 3.11). The QTL plot for PS (Figure 3.5) shows a peak location at 124 cM in a broad peak at the end of the chromosome plus a secondary sharp peak at 105 cM, for which there is slightly less statistical support, and which is coincident with the TBA and LS QTL. Despite the twin peak appearance of the PS QTL plot, a two-QTL model for LS was not significantly better than a one-QTL model.

These results show that the addition of markers in a region, where the density was low, can change results of a QTL analysis. Despite the absence of significant QTL

for OR and TN in this study, the earlier study (King *et al.*, 2003) detected a putative QTL at 49 cM for TN that reached the 5% chromosome-wide significance level. The differences in the results for QTL with effects on TN between the current and the earlier study may be due to the inclusion of additional markers enabling better tracking of chromosomal fragments through the pedigree. The additional markers are all distal to the previously reported TN QTL at 49 cM. One of the markers (SW268) close to this TN QTL was omitted from the current analyses as the genotypes did not pass the quality control checks. Although the current analyses suggest that the TN QTL found previously is probably a false-positive, other groups have reported QTL with effects on TN on SSC8 (Cassady *et al.*, 2001; Beeckmann *et al.*, 2003; Sato *et al.*, 2006; Bidanel *et al.*, 2008; Ding *et al.*, 2009).

The effects of the QTL found in this study were all negative dominant, i.e. the heterozygotes show inferior performance to both classes of homozygotes. In the first analysis, the additive effect was positive for TBA and LS and negative for PS. In the second analysis, all the QTL presented negative additive effects. Although the additive effects were not significant, the beneficial alleles at this QTL appear to be from the MS breed. This effect of the MS alleles would be consistent with previous observations describing the superior performance in LS in MS, through a higher level of ES for a given OR (Bidanel *et al.*, 1989; Haley & Lee, 1993).

For SSC8 numerous QTL for OR (Rathje *et al.*, 1997; Rohrer *et al.*, 1999; Wilkie *et al.*, 1999; Braunschweig *et al.*, 2001; Campbell *et al.*, 2003; Campbell *et al.*, 2008) and TN (Cassady *et al.*, 2001; Beeckmann *et al.*, 2003; Holl *et al.*, 2004; Sato *et al.*, 2006; Bidanel *et al.*, 2008; Ding *et al.*, 2009) have been found in previous studies. Also, a QTL for NSB have been reported recently in this chromosome by Li *et al.* (2009) in a White Duroc x Chinese Erhualian population. The reason for the discrepancy between studies is, as explained in the previous chapter, the different breeds used in the different analyses, with different genetic effects for the same traits.

The comparison of the linkage map used in the QTL analyses with the physical map, as deduced from the draft pig genome sequence (Sscrofa9), demonstrates that, with one exception, the linkage and sequence maps are co-linear. The order of SW2410 and HD, end markers at the beginning of the chromosome, are inverted relative to

one another on these maps. Since the end marker is only flanked on one side, it is difficult to check anomalies in the order of the end, and the penultimate marker on linkage maps. There are also uncertainties in draft genome sequences in this region. Finally, the human-pig comparative map does not assist with resolving this inconsistency, as there is no match in the human genome for the SW2410 sequence. SSC8 is a metacentric chromosome. The alignment of the recombination and sequence maps of SSC8 confirms that the frequency of recombination is greater at the telomeric ends of the chromosome than near the centromere.

The alignment of the linkage map used in the QTL analysis to the pig genome sequence allows the sequence to be searched for positional candidate genes for the QTL/trait. The draft pig genome sequence (Sscrofa9), which is available in the Ensembl genome browser, is incomplete. The sequences of some genes are completely missing from the draft genome assembly. Missing parts of some other genes stop the gene(s) from being recognised by the automated annotation systems. The Ensembl Biomart data-mining (http://www.ensembl.org/biomart/index.html) tools were used to identify genes in the QTL region (Appendix 2), between 90 and 120 Mbp. The casein genes (*CSN1S1, CSN2, CSN3*) shown at ca. 119 Mbp in this gene list indicates that there are some errors in the genome assembly, as these genes have been mapped by linkage analysis very close to the fibrinogen genes (*FGG*) (Archibald *et al.*, 1995). Chromosome 8 is stimated to be 120 Mbps in the Sscrofa9 assembly. These linkage map positions are supported by comparative genome mapping data; the human *CSN1S1* gene is located at 70.8 Mbp on HSA4 and this location has been shown to be homologous to the central part of SSC8 (Meyers *et al.*, 2005), rather than to the telomeric location suggested by the genome sequence.

Gene Ontology classifications of the positional candidate genes were also extracted using the Biomart data-mining tools. An examination of the Gene Ontology biological processes terms and the literature revealed a number of interesting genes from the list of >150 candidates: *TAC3R*, tachykinin receptor 3; *BMPR1B*, bone morphogenetic protein receptor, type 1B; *SPP1*, secreted phosphoprotein 1; *HPSE*, heparanase.

Bone morphogenetic proteins (*BMP*) are members of the transforming growth factor (TGF)-β family, with a role in bone formation during embryogenesis and fracture repair (Rosen & Thies, 1992). BMP have different receptors, from which Bone morphogenetic Protein Receptor 1B (*BMPR1B*) was considered as a candidate gene for reproductive traits, due to its location on SSC8 (Wang *et al.*, 2003). A single amino acid substitution in the ovine BMPR1B has been associated with the Booroola phenotype (FecB), which causes hyperprolificacy in Merino ewes through increases in OR (Davis *et al.*, 1987; Souza *et al.*, 2001; Wilson *et al.*, 2001). In sheep this mutation has been mapped to chromosome 6 in a region between *SPP1* and *EGF* (Montgomery *et al.*, 1994). These findings in sheep provided the original reasons to examine the homologous pig chromosome (SSC8) for QTL with effects on reproductive traits. QTL affecting OR have been mapped to this region (Rathje *et al.*, 1997; Campbell *et al.*, 2003), but no association was found between OR and *BMPR1B* (Wang *et al.*, 2003). In the previous analysis (King *et al.*, 2003), the *BMPR1B* gene was outside the confidence interval of the PS QTL, and no QTL for OR was detected. Moreover, the increased LS in MS has been attributed to an improvement in PS, rather than in OR (Haley & Lee, 1993). Kim *et al.* (2003) studied *BMPR1B* mRNA expression in endometrium, finding *BMPR1B* upregulated during the oestrous cycle when compared with early stages of pregnancy in gilts. Tomas *et al.* (2006) found a suggestive association between *BMPR1B* and both TBA and number weaned (NWEAN) in the first parity of an Iberian x MS population analysis. In more recent studies, an association between *BMPR1B* and preweaning survival has been suggested (Casellas *et al.*, 2008). Recently, in a microarray analysis *BMPR1B* was found differentially expressed between Chinese Taihu and LW pigs and there was a suggestive association of this gene with TBA (Sun *et al.*, 2011). In the present study, *BMPR1B* mapped between SW790 and SW61, where the peak for TBA and LS is located, suggesting *BMPR1B* as a candidate gene for these traits in this population.

Another gene of interest was *SPP1*, which was previously shown to be associated with LS in two different studies (van der Steen *et al.*, 1997; Korwin-Kossakowska *et al.*, 2001) and with LS and PS in the previous analysis in this population (King *et al.*, 2003). *SPP1* expression and regulation in reproductive tissues has been extensively

studied, not only in pigs but also in sheep (Johnson *et al.*, 1999a; Johnson *et al.*, 1999c; Garlow *et al.*, 2002; Kim *et al.*, 2003; Johnson *et al.*, 2003a; White *et al.*, 2005; Erikson *et al.*, 2009; Johnson *et al.*, 2009), mouse, rabbit, baboons, goats, and human (Nomura *et al.*, 1988; Waterhouse *et al.*, 1992; Fazleabas *et al.*, 1997; Johnson *et al.*, 1999b; Apparao *et al.*, 2001; Apparao *et al.*, 2003; Johnson *et al.*, 2003c; Joyce *et al.*, 2005; White *et al.*, 2006; Herington & Bany, 2007). For this reason, *SPP1* is a physiological and positional candidate gene, which is expressed in a variety of tissues, including the gravid uterus and placenta, where it has an important role to play in embryo implantation and maintenance of pregnancy. Moreover, SPP1 expression has been observed in ovine uterine (Johnson *et al.*, 2003b) as part of the focal adhesions between endometrial luminal epithelium (LE) and trophoblast (Tr), bound to integrins. In contrast to the earlier study (King *et al.*, 2003; King, 2003), the QTL for LS and PS were not co-located in the present study. However, *SPP1* is still a candidate gene due to its position under the peak for PS. Furthermore, given the location of SPP1 in reproductive tissues and the temporal regulation in expression during pregnancy, *SPP1* remains an important physiological candidate gene in this study.

Heparanase (HPSE) is an endoglycosidase that cleaves heparin sulphate side chains from Heparan sulphate proteoglycans (HSPGs) (Vlodavsky *et al.*, 1999), which have been found on the surface of the plasma membrane and in the ECM of various cell types, with a role in cell adhesion, migration, differentiation and proliferation (Wight *et al.*, 1992). Several groups have reported detection of HPSE activity in various tumour cells, platelets and placenta (Goshen *et al.*, 1996; Freeman & Parish, 1998; Vlodavsky *et al.*, 1999). Expression of HPSE has also been found in placenta in humans (Goshen *et al.*, 1996; Dempsey *et al.*, 2000; Haimov-Kochman *et al.*, 2002; Hasengaowa *et al.*, 2006), bovine (Kizaki *et al.*, 2001), mice (Revel *et al.*, 2005; D'Souza *et al.*, 2007), primate (D'Souza *et al.*, 2008), and in a recent study in pig (Miles *et al.*, 2009). As described previously, during the implantation and placentation of the embryo, an extensive remodelling of the ECM takes place, including angiogenesis, which relates with a function of HPSE during this period (Kizaki *et al.*, 2001).

Miles *et al.* (2009) mapped the porcine *HPSE* gene between marker SW1980 at 126.1 cM and S0178 at 127.7 cM based on MARC-USDA marker positions, and located *HPSE* mRNA in the pig placenta throughout gestation. The *HPSE* gene is located under the peak of the QTL for LS and PS mapped in the previous analysis (King *et al.*, 2003), and under the peak of the PS QTL in the present study. Miles *et al.* (2009) hypothesised that the reported differences in the morphology of the folded luminal epithelium and placenta between the smallest and largest foetus in a litter (Vallet & Freking, 2007) could be measured as variations in *HSPE* mRNA expression. Therefore, Miles *et al.* (2009) suggested *HPSE* as a candidate gene for LS and PS due to its role in placenta, and its chromosomal location. Due to the mapping of *HPSE* to the QTL region revealed in the previous analysis performed on this chromosome for this population, in the second fine mapping performed in this study, a microsatellite marker was designed from a BAC clone sequence in the region of this gene (CH242-443f10). As a result of this genotyping, this marker was located nearly 10 cM distal to *SPP1* in the linkage map, and 5 cM apart from the peak of the PS QTL (Figure 3.5), but still under the broad plateau of this PS QTL. Therefore, *HPSE* remain a candidate gene for reproductive traits in this population and suggest the need of further analyses.

TAC3 encodes Neurokinin B (NKB), which is a member of a family of neuropeptides called the tachykinins, that was suggested to play a role in trophoblast invasion occurring during implantation (Page *et al.*, 2000). The expression of this tachykinin has been found to be higher in placental tissue than in any other organ or tissues in the human body (Page *et al.*, 2000; Page *et al.*, 2006). During pregnancy, the activation of tachykinin receptor 3 (NK3R encoded by *TAC3R*) by NKB reduces the large blood flow through the liver to satisfy the need of the uterus and placenta (Pinto *et al.*, 1997). The presence of *TAC3* mRNA has been reported in human and rat placenta (Page *et al.*, 2000), in rat uterus (Barr *et al.*, 1991; Cintado *et al.*, 2001; Candenas *et al.*, 2001), and in mouse placenta, uterus and oocytes (Pinto *et al.*, 2001; Pinto *et al.*, 2009). The level of expression of NKB and NK_3R in the female reproductive system are maximal around implantation (Page *et al.*, 2000), and elevated levels of NKB have been detected in pre-eclampsia in humans (Page, 2010). The level of NK_3R in the rat uterus has also been shown to vary during pregnancy,

with very low levels during late pregnancy compared to high expression during the early stages (Candenas *et al.*, 2001), the expression of both, NKB and NK_3R, is under oestrogenic control in rats and mice (Pinto *et al.*, 1997; Pinto *et al.*, 2009). In view of the results obtained in the analysis of this tachykinin and its receptor in mouse, rat and human the study of this gene in pig is warranted, due to its position in the genome and the function showed in other organisms.

Comparative mapping was used to assess the similarities between SSC8 and homologous regions in the human genome initially as a means of identify comparative positional candidate genes and also as a check on the linkage map. Both these functions were largely replaced by the draft pig genome sequence when it became available. HSA4 shares extensive homology with SSC8, as well as with SSC15 and SSC17, but gene order differs between HSA4 and SSC8. In humans the genes *IBSP*, *SPP1*, *DSPP*, *AMBN*, and *BMP3* are all associated with mineralised tissues, and all map close together on HSA4. But in pig, there is an inversion in the homologous region in SSC8, so these genes are not together (Jiang *et al.*, 2002a). The greater knowledge of the gene content of the homologous human chromosomal region on HSA4 can be used to identify further genes which may represent (comparative) positional candidates for the SSC8 QTL.

The results presented here represent and confirm the importance of SSC8 in reproductive traits in pig. The comparative mapping, both with the human and pig sequence, reveal the large number of genes present in a QTL region. As mentioned previously, increasing the number of trait recorded and genotyped animals are the most effective means of improving the resolution of QTL and trait gene mapping studies and confidence in the results. However, the fine mapping of the SSC8 QTL region mapped in the previous study in the same population has changed the understanding of the QTL locations and the positional candidate genes for these QTL effects. With the advance in the pig genome sequence, the number of genes mapped is increasing. Thus, the number of positional candidate genes to investigate is large. As a result, in this study several genes with possible functions in reproduction were mapped to the QTL, making these genes positional and physiological candidates for the traits of interest. For a confirmation, and in order to use these genes for MAS,

they should be mapped as candidate genes in other populations, and extensive functional analyses carried out, to confirm the possible contribution of these genes and their potential to contribute to improvements in reproductive performance. In this study, *SPP1* was the candidate gene chosen for further analyses and the results of the characterisation study are presented in Chapter 6.

Chapter 4
Genome-Wide Association Study of genetic variation in reproductive traits

4.1. Introduction

In the previous two chapters a QTL scan for reproductive traits is described. This scan was performed in an experimental population comprising The Roslin Large White – Meishan structured pedigrees and using a linkage-based approach with DNA markers, mainly microsatellites. As a result, a number of QTL and candidate genes were identified. As noted earlier, the number of trait recorded and genotyped animals determine the power of such genetic studies. The cost of establishing experimental populations is a significant limitation on such studies.

Genome-Wide Association Studies (GWAS) which exploit population-wide linkage disequilibrium offer a powerful alternative approach, as demonstrated in studies of a wide-range of traits in humans. In GWAS studies, several hundred to several thousand individuals for whom the trait of interest has been recorded are each genotyped for several thousand Single Nucleotide Polymorphisms (SNPs). The resulting data are analysed for evidence of associations between variation in the trait of interest and SNP genotypes. In this study, genotypic data were obtained using the Illumina Porcine SNP60 BeadChip (Illumina, San Diego, CA) (Ramos *et al.*, 2009) on samples from a commercial pig population. The resulting genotypes, together with the phenotypic data collected and the information available on the position of the SNPs, were used in a genome-wide association study (GWAS) to identify SNPs or genomic regions associated with reproductive traits. In total 404 animals were genotyped, and a whole genome association analysis was performed using the R package GenABEL (Aulchenko *et al.*, 2007b). Associations were performed using the residual values from a linear mixed animal model analysis of each trait (Aulchenko *et al.*, 2007a). The SNP effects were then re-estimated from linear mixed model analyses of the data in which the significant SNPs were fitted, individually, as additional fixed effects.

4.1.1. SNPs and previous studies

SNPs are the most common source of variation in vertebrate genomes. Thus, they are valuable tools for linkage and association studies, which require a large number of genetic markers as well as a large number of animals with quality phenotypic recording and DNA. The porcine SNP60 Beadchip was developed through the efforts of the International Swine Genome Sequencing Consortium (http://piggenome.org/). This high density swine SNP chip contains probes for a total of 64,232 SNPs which had been identified in commercial European and US breeds (Duroc, Landrace, Large White, Pietran and Wild boar) (Ramos *et al.*, 2009).

Recently the number of studies using this chip has increased, not only for reproductive traits (Onteru *et al.*, 2011; Onteru *et al.*, 2012) but also for production traits (Duijvesteijn *et al.*, 2010; Fan *et al.*, 2011; Grindflek *et al.*, 2011; Ponsuksili *et al.*, 2011; Ramos *et al.*, 2011). Onteru *et al.* (2011) studied lifetime reproductive traits of 818 gilts from a LW and a LW x LR population over 9 parities, and found that a total of 59 regions were associated with the traits analysed. On closer inspection, only eight of these regions were from previously reported QTL regions, and of these eight QTL regions only three were associated with reproductive traits. The same group (Onteru *et al.*, 2012) reported results from an association study for reproductive traits (LS, TBA, NSB, NMUM and GL) in 683 females pigs over the first three parities from the same population used in their previous studies (Fan *et al.*, 2009; Onteru *et al.*, 2011). The association analysis performed separately for each parity identified different genes affecting each trait in the different parities that had not been identified previously as candidate genes for these traits.

4.2. Materials and Methods

4.2.1. Population information and phenotypic traits recorded

The data were collected from 4,378 litters from 1,019 sows in a commercial multi-line multiplication herd that had been recorded continuously for reproductive traits.

During this period, there were two disease outbreaks in the population. However for this study only healthy animals were used. The herd, located in China, was composed among others of LR, LW, Duroc, Pietrain, MS, and synthetic lines (Duroc x White, LW x White, White x MS, White x Duroc). Full pedigree information was available for every animal going back five generations, giving a total of 4,104 animals in the pedigree. DNA was extracted by PIC/Genus from all animals with phenotypic records, and all SNPs genotyping was done at the Wellcome Trust Clinical Research Facility of the Western General Hospital, University of Edinburgh. In a previous study (Lewis *et al.*, 2009b), statistical analysis comparing performance differences between sows of different parities and lines was performed.

For each litter, recorded data included sow identity, dam, sire, farrowing date, service date, services (total times females came into oestrus before holding to service), matings per conception (total number of inseminations until conception), gestation length (GL), litter size (LS), number of piglets fostered off, number of piglets fostered on, total born alive (TBA), total piglets born dead (TBD), number of mummified piglets (NMUM), number of stillborn piglets (NSB), lactation length (LL) and number weaned (NWEAN).

4.2.2. Exploratory statistics and parameter calculations

Exploratory analysis of the data was performed with SAS (SAS Institute Inc., Cary, NC, 2006), in order to calculate the mean and their standard error, and to test whether the data met the expectations of a normal distribution. For non-normally distributed traits, a log transformation was performed; with variables that contained zero values being transformed using log (trait+1) (Lewis *et al.*, 2009a). Genetic parameters, i.e. heritabilities, and their standard errors, were estimated using the ASREML package (Gilmour *et al.*, 2009), fitting an animal model including all known pedigree relationships and fitting sow line and parity as fixed effect and animal id as random effect.

4.2.3. Quality control (QC)

The SNP genotype data included information for 62,163 SNPs on 404 animals. These data were subjected to quality control (QC) measures using the GenABEL program. The data were checked for marker call rate (<0.95) and minor allele frequency (MAF) (<0.01), which allowed the identification and removal of SNPs with MAF less than 0.01 or for which the SNP genotyped could not be determined in >5% of the samples. The quality checks on the samples involved a) identifying and removing samples for which <95% of the SNP assays yielded a genotype, b) identifying and removing samples for which the SSCX SNP genotypes indicated that the DNA was from a male pig; and c) identifying and removing samples for which the SNP genotypes were >95% identical across all markers (i.e. Identical-By-State (IBS)). Due to the population admixture (sows from many differing lines), there was no expectation of Hardy-Weinberg equilibrium and therefore no filter for this was used in the QC. In total, 3,669 markers were excluded because of low call rate (<95%), and 4,994 were excluded because of low MAF. Regarding the animals, 48 were excluded because of low call rates, 13 samples/animals were excluded as they were probably males, and 3 animals were excluded because high IBS value. After editing, the data included 53,501 SNPs and 340 animals (Table 4.1).

	SNP markers		**Sample**	
	Removed	**Passed**	**Removed**	**Passed**
Initially		62,163		404
Call rate <0.1	3,657	58,506	48	356
Female/male		58,506	13	343
MAF <1% (Run 1)	4,988		-	
Call rate <95%	12		0	
IBS>=0.95		53,507	3	340
MAF <1% (Run 2)	6	53,501	-	-
Final		**53,501**		**340**

Table 4.1 Quality control summary, indicating number of SNPs and samples removed and the ones that passed the control.

4.2.4. Genome-wide SNP association analysis

The SNP association analysis was performed separately for each trait (raw and transformed traits), using the GenABEL (Aulchenko *et al.*, 2007b) package in R with a GRAMMAR approach (Genome-wide Rapid Association using Mixed Model and Regression) (Aulchenko *et al.*, 2007a). The analysis was performed as described by Aulchenko *et al.* (2007a), i.e. regressing residuals obtained from a mixed model analysis of each trait on the SNP genotypes. The residuals of the original phenotypes were obtained from a mixed model fitted using the ASReml software fitting the fixed effects of sow line (9 levels) and parity (10 levels), and the sow fitted as a random effect using the full pedigree (4,104 sows in the full pedigree). Therefore, the residuals were adjusted for family and environmental effects using a polygenic model. Since there was more than one litter record per sow, the average of the residuals for each sow for each trait was calculated, assuming that traits on different litters from the same sow were repeated measurements of the same trait. The positions of the SNPs used in this analysis were from the 2009 draft genome sequence (Build9, Sscrofa9) as annotation of this assembly is available in the Ensembl genome browser and associated database.

A single SNP trait-association analysis was performed, without permutation, and fitting the principal components that account for the structure of the data. The method used was the one described by Price *et al.* (2006) (EIGENSTRAT), which makes use of principal components of the genomic kinship matrix to adjust both phenotypes and genotypes to account for the stratification. First, the kinship matrix was generated in order to check for genetic stratification of the population. Second, the principal components of genetic variation were calculated using the kinship matrix.

The first ten principal components were extracted from the analysis and added to the phenotype file, only for the samples that passed the QC. Finally, the **qtscore** function using an additive model was used, including the best two principal components, to identify the genome-wide significant SNPs for services, matings per conception, GL,

LS, TBA, TBD, NMUM, NSB, LL and NWEAN, and the transformed traits for mating, services, TBD, NMUM and NSB. The threshold for confirmation of significant results was set at a *p*-value of less than 0.05 divided by the actual number of SNPs tested (9.34×10^{-7}). In this analysis, permutations were not run. For this reason, uncorrected *p*-values of $P < 5 \times 10^{-5}$, as a minimum, were accepted and considered to have significant association with the trait. These significant SNPs found in this analysis, were further explored.

4.2.5. Exploration of SNP effects

SNPs that were found to be significant in the genome-wide association analysis were further explored in mixed model association analyses, in order to estimate their effects. The model used was an extension of that described in the previous section. Briefly: line, parity and the SNP genotype were fitted as fixed effects in an animal mixed model including sow as a random effect and accounting for the full pedigree. Different litters from the same sow were treated as repeated measurements.

Predicted trait values for each genotypic class of each SNP were obtained from the ASReml analyses. The predicted trait values were used to estimate additive and dominance effects on traits for each SNP, and the proportion of additive genetic variance (V_A) for each trait accounted for by the SNPs. The equations used were: additive effect, $a = (AA - GG)/2$; dominance effect, $d = AG - [(AA + GG)/2]$; and % V_A due to the SNP $= [2pq(a + d(q - p))^2]/V_A$ where AA, AG and GG were the predicted trait values for each genotype class, p and q were the allelic frequencies at the SNP locus, and V_A was the additive genetic variance of the trait obtained from an animal model analysis ignoring the SNP effects.

4.2.6. Candidate genes

The regions, where significant associations were found in the baseline dataset, were explored for positional candidate genes within 3 Mbp either side of significant SNPs (Du *et al.*, 2007).

4.3. Results

4.3.1. Data summary and heritability of the traits

A full description of the data is found in Lewis *et al.* (2009b). Summary statistics for all observed reproductive traits are shown in Table 4.2.

	Mean	SD	min	Max
LS	11.050	2.983	0	23
TBA	10.235	3.002	0	21
TBD	0.815	1.314	0	23
Trans1 TBD	0.433	0.527	0	3.17
NMUM	0.216	0.659	0	12
Trans1 NMUM	0.125	0.320	0	2.56
NSB	0.599	1.062	0	23
Trans1 NSB	0.339	0.467	0	3.17
NWEAN	9.450	2.061	0	24
GL (days)	115.57	1.638	108	126
LL (days)	21.82	4.577	0	61
Services	1.079	0.309	1	5
Trans Services	0.052	0.194	0	1.60
Mating	1.910	0.765	1	10
Trans Mating	0.579	0.366	0	2.30

Table 4.2 Table showing descriptive statistics (mean, standard deviation (SD) and range min-max) for all the traits and the transformed traits. Trans, log (trait) transformation; Trans1, log (trait + 1) transformation.

4.3.2. Heritabilities

Heritability estimates and standard errors for all traits analysed using a linear model are presented in Table 4.3. Heritability estimates were generally low for all traits but GL.

	Heritability	SE
LS	0.202	0.018
TBA	0.174	0.018
TBD	0.046	0.012
Trans1 TBD	0.051	0.012
NMUM	0.029	0.011
Trans1 NMUM	0.029	0.011
NSB	0.042	0.011
Trans1 NSB	0.045	0.012
NWEAN	0.024	0.011
GL (days)	0.458	0.019
LL (days)	0.067	0.014
Services	0.124	0.182
Trans Services	0.038	0.012
Mating	0.063	0.013
Trans Mating	0.068	0.013

Table 4.3 Heritability (h^2) estimates and standard error for reproductive traits. Trans, log (trait) transformation; Trans1, log (trait + 1) transformation.

4.3.3. SNP association analysis

The SNPs showing significant association with the traits analysed in the GWAA are presented in Table 4.4. As permutations were not performed in this analysis, the *p*-values of the SNPs presented here, corrected for lambda inflation factor by GenABEL, are under the *p*-value calculated for P<0.05 with Bonferroni corrections. In the Table 4.4, the position of each SNP in the pig genome is indicated, with chromosome number (SSC) and position in basepairs (bp) for Sscorfa9 which is the genome assembly currently available through the Ensembl genome browser (http://www.ensembl.org/Sus_scrofa/Info/Index).

Trait	SNP	SSC	Position (bp)	Pc1df
LS	ALGA0103270	8	8,604,503	2.95E-06
	ASGA0099069	0	0*	9.60E-06
TBA	ALGA0103270	8	8,604,503	8.14E-06
	ASGA0099069	0	0*	1.78E-05
TBD	MARC0052517	17	16,793,187	6.89E-06
NMUM	ALGA0084499	15	24,916,712	4.89E-07
	ASGA0019012	4	20,983,748	2.26E-06
	ALGA0121141	6	82,292,678	2.46E-06
Trans1 NMUM	H3GA0055446	2	15,284,268	2.65E-05
	ALGA0084780	15	29,597,315	2.99E-05
	ALGA0084499	15	24,916,712	5.32E-05
Trans1 NSB	BGIS0004826	14	57,281,024	4.27E-05
Gestation length (days)	ALGA0021148	3	107,708,164	2.70E-05
	H3GA0055694	15	29,058,423	4.78E-05
Lactation Length (days)	ALGA0004694	1	87,542,756	5.50E-07
	MARC0015922	1	86,751,931	2.61E-06
Trans Services	MARC0036115	3	63,400,919	1.71E-14
	MARC0082152	3	67,987,672	3.23E-08
	INRA0015162	4	82,561,154	3.62E-06
Mating	MARC0092197	9	119,688,265	1.98E-05
	ASGA0047195	10	24,653,624	2.30E-05

Table 4.4 List of SNPs associated with reproductive traits. The Table shows the position of the SNPs (SSC and bp), and P-value with 1df, calculated in the principal components analysis. Pc1df, P-values adjusted for genomic control; Trans, log (trait) transformation; Trans1, log (trait + 1) transformation. * The sequence of the ASGA0099069 SNP is not present in the Sscrofa9 assembly. A BLAST search of the most recent assembly (Sscrofa10.2) mapped this SNP to SSC8 at 11,003,426 bp close to the ALGA0103270 SNP which maps to SSC8 at 10,993,631 bp in Sscrofa10.2.

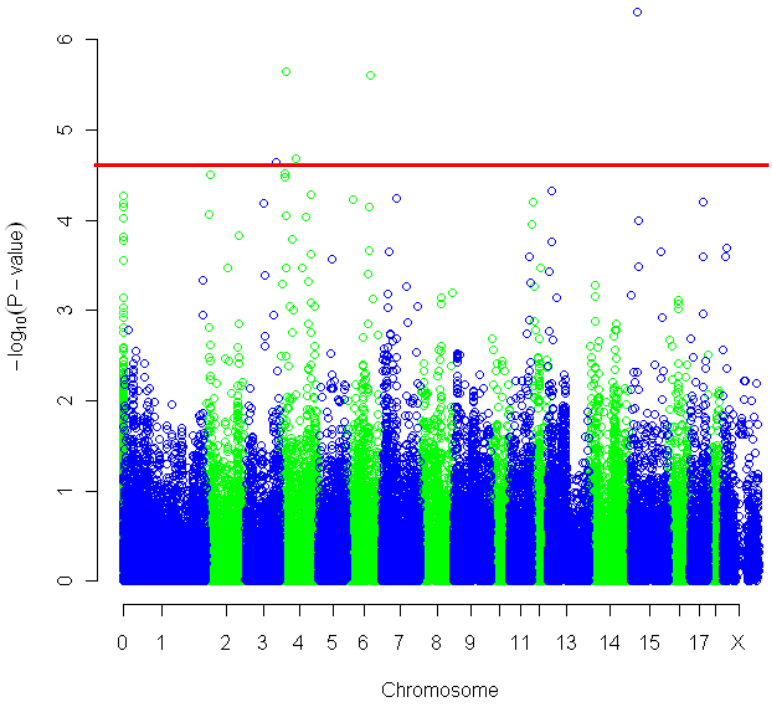

Figure 4.1 Manhattan plot displaying the results (-log10 of P-value) for total mummified piglets trait analysis. Red line indicate P=5 x 10^{-5}.

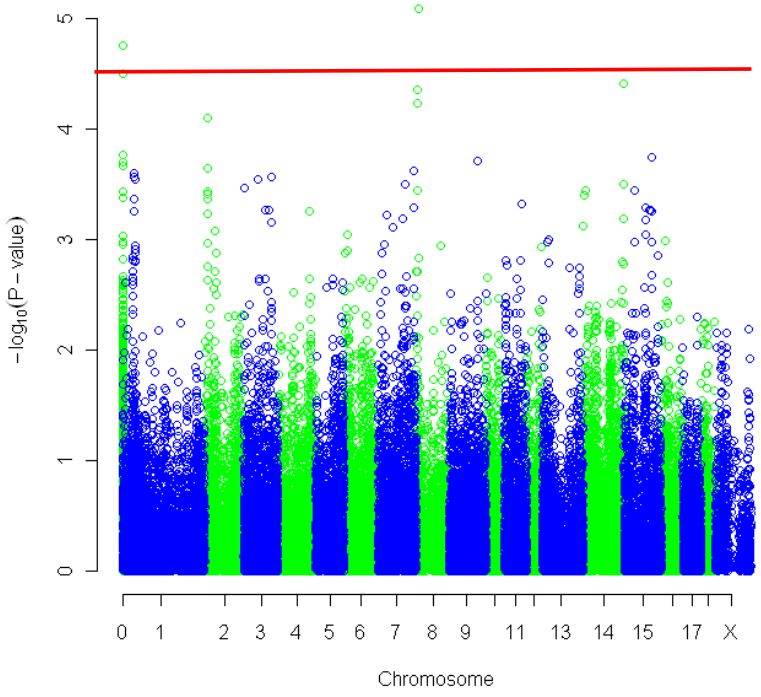

Figure 4.2 Manhattan plot displaying the results (-log10 of p-value) for total number of piglets born alive. Red line indicate P=5 x 10^{-5}.

Figures 4.1 and 4.2 represent the results from the GWAA for two of the traits where further analysis was performed for the significant SNPs.

4.3.4. Significant SNP effects and candidate genes

Out of 21 SNPs listed in **Error! Reference source not found.**, 17 were further analysed to estimate their effects. The results of additive and dominance SNP effects, and the genetic variation explained by the SNP are presented in Table 4.5. Positional candidate genes for these regions are listed in Appendix 3. Relevance for some of these genes is discussed in the discussion section.

SNP	Trait	P-value GenABEL	P-value SNP model	Additive effect (±se)	Dominance (±se)	h2q	%VA explained by SNP	Allele freq	
								p	*q*
ALGA0103270	**LS**	3.17E-06	0.005	0.574 (0.345)	0.005 (0.259)	0.01	0.07	0.76	0.24
ASGA0099069	**LS**	1.14E-05	0.008	0.567 (0.339)	0.012 (0.253)	0.01	0.07	0.76	0.24
ALGA0103270	**TBA**	8.62E-06	0.005	0.548 (0.303)	0.088 (0.253)	0.01	0.06	0.76	0.24
ASGA0099069	**TBA**	2.24E-05	0.016	0.524 (0.339)	0.073 (0.248)	0.01	0.06	0.76	0.24
MARC0052517	**TBD**	6.89E-06	0.001	-0.435				0.99	0.01
ALGA0084499	**NMUM**	4.89E-07	0.001	-0.774 (0.148)	-0.475 (0.181)	0.05	0.32	0.98	0.02
ASGA0019012	**NMUM**	2.26E-06	0.001	-0.723 (0.085)	-0.392 (0.127)	0.02	0.58	0.97	0.03
ALGA0121141	**NMUM**	3.07E-06	0.001	-0.776 (0.044)	-0.601 (0.157)	0.00	0.15	0.98	0.02
ALGA0084499	**Trans1 NMUM**	5.32E-05	0.001	-0.264 (0.028)	-0.123 (0.102)	0.01	0.27	0.98	0.02
H3GA0055446	**Trans1 NMUM**	2.65E-05	0.007	-0.062 (0.030)	-0.010 (0.030)	0.01	0.30	0.82	0.18
ALGA0084780	**Trans1 NMUM**	2.99E-05	0.001	-0.008 (0.040)	0.102 (0.031)	0.01	0.53	0.83	0.17
BGIS0004826	**Trans1 NSB**	1.60E-05	0.001	-0.073 (0.052)	0.017 (0.026)	0.00	0.06	0.54	0.46
ALGA0004694	**LL (days)**	5.50E-07	0.001	-0.943	-	-	-	0.94	0.06
MARC0015922	**LL (days)**	2.61E-06	0.002	-0.962				0.95	0.05
MARC0036115	**Trans Services**	1.71E-14	0.003	-0.071				0.99	0.01
INRA0015162	**Trans Services**	3.62E-06	0.146	-				0.99	0.01
MARC0082152	**Trans Services**	3.23E-08	0.001	-0.175 (0.032)	0.006 (0.080)	0.03	0.73	0.76	0.24

Table 4.5 Table showing significant SNPs verified with ASReml. The Table shows The *P*-values from GenABEL analysis together with *p*-values from ASReml analysis for each SNP and trait analysed. The additive and dominance effects, and the proportion of genetic variance explained by each significant SNP and the allele frequencies calculated, are summarised in this Table. Trans, log (trait) transformation; Trans1, log (trait + 1) transformation.

4.4. Discussion

In the present study, a GWAS using the Porcine SNP60 BeadChip was performed using a GRAMMAR approach for different reproductive traits recorded in a commercial population. As this population included animals from different pig breeds it was necessary to account for the stratification of the population in the analysis using the principal components of genetic variation. The genome-wide association analysis identified associations for a total of 17 SNPs from eight different chromosomes with 6 different traits. Visual inspection of the Manhattan plots (see Figures 4.1 and 4.2), where p-values for all the SNPs are represented by position in the genome, shows that there are additional SNPs in the same regions as the significant SNPs which are evident above the background but which do not meet the significance threshold. However, due to the exploratory nature of the analysis performed in this study, only the most significant SNPs were analysed for their effects for each trait. Thus, additive and dominance effect and the proportion of the variance explained by the SNP were calculated for the significant SNPs. For some of the significant SNPs, homozygote animals for one of the alleles were missing. Therefore for these SNPs only the additive effect was calculated. Besides, some of the SNPs have very low alleles frequencies, due to the use of a low MAF in the quality control which is less stringent and results in false positive, but allows the inspection of more SNPs as well as more candidate genes with caution.

In the present study, two SNPs (ALGA0103270, ASGA0099069) on SSC8 were associated with LS and TBA trait. The position of these SNPs at ~11 Mbp on the latest genome assembly (Sscrofa10.2) (Table 4.4) indicated no relation of this SNP with the QTL mapped in the previous chapter for these traits. The QTL with effects on these traits identified in the linkage analyses and described in Chapters 2 and 3 are located towards the telomeric end of SSC8q in contrast to the SNP effects observed here towards the telomeric end of SSC8p – the other end of this metacentric chromosome. However in a recent study, Onteru *et al.* (2012) found an association of a region on SSC8 at position 15.96 Mbp with LS in a similar study. The position of these SNPs is not the same but it could be considered to be in the same region. In this

region of SSC8, numerous QTL for OR have been mapped in diverse populations (Rathje *et al.*, 1997; Rohrer *et al.*, 1999; Wilkie *et al.*, 1999; Braunschweig *et al.*, 2001; Jiang *et al.*, 2001; Campbell *et al.*, 2003). As described previously, LS is a composite trait and OR is one of the traits affecting it.

For the NMUM trait, both the raw phenotypic data and log transformed phenotypic data were tested for associations. Associations were observed with two SNPs on SSC15, of which one (ALGA0084499) was significantly associated with both the untransformed and transformed trait, and the other (ALGA0084780) has a significant negative additive effect. In a previous study, Onteru *et al.* (2012) found an association between this region of SSC15 and both, TNB and NSB traits. SNPs on SSC2, SSC4 and SSC6 were also found to be associated with the NMUM trait. Holl *et al.* (2004) have reported QTL on SSC2, SSC6 and SSC12 with effects on this trait (NMUM). Two of these QTL were mapped to different positions on SSC2p with peaks at 6 and 29 cM on the linkage map used (Cassady *et al.*, 2001; Holl *et al.*, 2004). The genetic markers SW1514 and SW1515 at 0 and 23 cM respectively on this linkage map (Cassady *et al.*, 2001), map to 2.12 and 14.87 Mbp on the Sscrofa9 genome assembly. Therefore, these results from Holl *et al.* (2004) provide support for the association between SNP H3GA0055446 and NMUM (transformed) observed in this study. Holl *et al.* (2004) also reported four QTL on SSC6 with effects on NMUM with peaks at 64, 81, 165 and 191 cM. Alignment of the genetic markers from the SSC6 linkage map used by Holl *et al.* (2004) with the genome sequence suggests that the ALGA0121141 SNP maps between 123 and 153 cM. As Holl *et al.* (2004) only report peak QTL positions and do not present the associated QTL plots, it is difficult to assess the overlap between the results of this study and their results. Onteru and colleagues report associations between SNPs on SSC6 and NMUM in a similar analysis to the one presented here, including an association of NMUM in parity 3 with MARC0043661-MARC0105315 which map to ~80 Mbp on SSC6 close to the ALGA0121141 SNP at 82 Mbp reported here (Onteru *et al.*, 2012). The importance of the reduction in ES and PS as a mean to increase LS have been discussed in previous Chapters, and for this reason NMUM is an important trait for

the improvement in reproduction. The SNPs reported here are potentially useful as they explain a high proportion of the genetic variance in the trait.

Another trait for which few QTL have been mapped is NSB. In this study, a single SNP (BGIS0004826) was identified on SSC14 at position 57,281,024, which was significantly associated with the transformed NSB trait. Two QTL with effects on NSB have been mapped to SSC14 by others; at 28 cM, corresponding approximately to 17-28 Mbp (Tribout *et al.*, 2008), and at 104 cM which represented the end of the linkage map used (Holl *et al.*, 2004). SNPs on SSC14 associated with NSB have also been reported at ~36 and 145 Mbp in parity 1 and at ~94 - 96 Mbp in parity 2 (Onteru *et al.*, 2012).

The study of Onteru and colleagues (2012) also provides some support for the SNP associations for GL observed in this study (Table 4.4). These authors report associations between SNPs on SSC3 at ~109 Mbp and SSC15 at ~20 Mbp and GL in gestation 2 and 1 respectively. Although a QTL for GL has been mapped to SSC15 by Wilkie *et al.* (1999) it is located at the other end of the chromosome from the SNP association reported here.

The Ensembl Biomart data mining tool (http://www.ensembl.org/biomart/index.html) was used to search a region of approximately 3 Mbp on both sides of the most significant SNPs listed in Table 4.4. The position of genes in these regions together with gene symbols, gene names and Gene Ontology process terms were exported and inspected for genes with functions relevant to the trait of interest. As the Sscrofa9 genome assembly available in Ensembl is incomplete these searches are only a first attempt to identify candidate genes.

The list of positional candidate genes for the regions surrounding the SNPs with significant associations with reproductive traits as observed in this study is both long and incomplete (Appendix 3). The Sscrofa9 genome assembly, from which these positional candidate genes were identified, is incomplete and therefore some genes are missing as the corresponding sequence is missing. Other genes are missing as the fragmented nature of the genome assembly stops the automated annotation tools

identifying their partial presence. The positional candidate genes listed are limited to protein-coding genes.

The positional candidate genes include several which are involved in multiple embryonic and developmental processes including *LRP2*, *AMBRA1*, *ALX4*, *EXT1*, *EN1*, *GLI2*, *EPB41L5* and *TBCE*. Variation in the expression of these genes or in the encoded proteins would be expected to have important consequences for the developing embryo. Therefore associations with traits such as NSB and NMUM pigs might be expected, but the associations observed are with the mother's genotype not those of the developing embryos.

Of the positional candidate genes on SSC8 for the association with TBA and LS, BST1 (bone stromal cell antigen 1, also known as CD157) facilitates pre-B cell growth and has been identified as a risk factor for Parkinson's Disease. According to the mouse gene expression data available on the BioGPS web site (http://biogps.org), Bst1 is expressed in mouse placenta. It has also been shown to interact with integrins (Lo Buono *et al.*, 2011). Of the other candidate genes in this region FGFBP1 is abundantly expressed in mouse umbilical cord and PROM1 abundantly in mouse uterus.

The NCOA1 gene, which is listed as a positional candidate gene for the GL association, has been considered previously as a candidate gene for reproductive traits (Melville *et al.*, 2002). The nuclear receptor coactivator 1 (NCOA1), also known as steroid receptor coactivator (SRC1), is a member of the nuclear receptor coactivator family. The steroid receptors bind steroid hormones such as oestrogens, progestins, androgens, glucocorticoids and mineralocorticoids. The NCOA1 protein enhances activity of the ESR1 receptor that, in turn, stimulates the transcription of specific oestrogen-responsive genes and mediates subsequent physiological responses. Melville *et al.* (2002) argued that on this basis the NCOA1 gene could be considered as a physiological candidate gene for prolificacy traits in pigs but found no significant associations with OR, LS or TBA in The Roslin LW-MS crosses. These authors did not test for associations with GL.

The SNP associations presented in this study are few but significant. As mentioned previously, the analysis presented here is only an exploratory analysis. The results presented demonstrate the potential of these data and indicate the need for further analysis. The data from different lines and the large number of parities present an opportunity for a more complex analysis to the one presented here. The different breeds need to be inspected in order to see where the different alleles come from. Treating the data from the different parities as separate traits would be worthwhile as it can be argued that the sows are in different physiological states in each successive pregnancy. As shown by Onteru *et al.* (2012) different SNPs can be associated with a trait in different parities. The variation in the number of litters per sow causes a more accurate measure of litter traits in those sows with more parities and thus, the data from different sows are not equivalent. This study could also benefit from a larger number of animals genotyped, an important factor affecting the significance of the results.

The results presented here failed to confirm any of the QTL with effects on reproductive traits identified in the linkage analyses of the LW-MS crosses. A possible explanation for these results is the difference between both studies, QTL and GWAS. A difficulty with GWAS studies in which tens of thousands of SNPs are used resulting in thousands of tests being performed in the analyses is setting the appropriate threshold for acceptance of results. For the *p*-value, traditional methods, such as Bonferroni correction, can result in too strict cut off points while relaxing the correction will increase the number of false-positive accepted. For this reason, independent studies are important for the confirmation of results.

Chapter 5
Validation of preservation and fixation methods for reproductive tissues and selection of internal control genes for quantitative studies

5.1. Introduction

The chromosomal regions defined in QTL mapping studies, as described here, typically harbour many positional candidate genes. Characterisation of positional candidate genes, in terms of levels and patterns of expression at the transcript and protein levels, may not only help identify the gene responsible for the QTL effects but also explain the effects. By assaying the level of expression by qPCR and the patterns of expression at the mRNA and protein level by *in situ* hybridisation (ISH) and immunohistochemistry (IHC), respectively, it is possible to address the following questions. Is the positional candidate gene expressed in a relevant tissue? Do the expression levels or patterns of expression of the candidate gene vary between individual pigs? For example, the *SPP1* gene which lies under the QTL for prenatal survival (PS) is expressed in reproductive tissues including placenta and endometrium with various patterns of expression at the different stages of pregnancy (Johnson *et al.*, 2009).

Experiments to characterise candidate genes for reproductive traits in pigs present a number of challenges. Pigs, including purebred individuals, are outbred and genetically heterogeneous in contrast to inbred strains of laboratory animals, such as mice and rats which are genetically homogeneous. It is desirable, therefore, to carry out as many of the different analytical methods as possible on each animal in order to compare, for example, the quantity and distribution of mRNA, or the distributions of mRNA and the encoded proteins. Pigs, especially pregnant sows, are large animals. Sacrificing and dissecting an adult pig in order to access the tissues of interest can take significantly longer than the corresponding procedures in laboratory animals. Pigs are also more expensive to purchase and maintain than laboratory animals. Pigs reach sexual maturity later than mice or rats, and have longer gestation intervals, exacerbating the maintenance costs and limiting the ease and speed with which studies of reproduction can be repeated. For all these reasons, it was critical that the methods for isolating and preserving samples were optimised before starting the experiments on the animals of interest.

5.1.1. Preservation of tissues

The isolation of high quality intact RNA is challenging, as RNA is readily degraded during the collection of tissues, due to the endogenous RNases present in the tissues at the moment of dissection. Degradation can also be caused during fixation, to different degrees depending on the fixative (Srinivasan *et al.*, 2002; Cox *et al.*, 2008), and during RNA extraction (Hamatani *et al.*, 2006). The structure of the tissue and its molecular characteristics can also suffer changes during the preservation process (Srinivasan *et al.*, 2002), affecting not only the RNA, but also the histology and protein profile of the tissue. The need to isolate adequate quality RNA for quantitative and functional studies required the search for an effective preservation method, with minimal mRNA and protein degradation.

For pathological assessment, paraffin-embedding is routinely used because of the ease of handling tissues and subsequent staining. Treating the tissue with fixatives such as Paraformaldehyde (PFA), Bouin's, and alcohol-based methods such as Carnoy's and Methacarn prior to embedding is an efficient way of preserving tissue structure. But some of these methods are not optimal for subsequent isolation of high quality RNA. Other methods, such as snap freezing in liquid nitrogen or preserving solutions containing organic solvents and denaturing agents (e.g. phenol), provide better RNA quality for functional genomics experiments but destroy tissue integrity, and consequently are suboptimal for tissue histology (Florell *et al.*, 2001). The perfect solution would preserve not only the tissue integrity but also the RNA quality.

Methacarn, an alcohol-based fixative (Puchtler *et al.*, 1970), was considered as a potential fixative based on previous studies demonstrating the preservation of RNA, protein and tissue structure (Puchtler *et al.*, 1970; Mitchell *et al.*, 1985; Tyrrell *et al.*, 1995; Shibutani *et al.*, 2000; Uneyama *et al.*, 2002; Takagi *et al.*, 2004; Lee *et al.*, 2006; Cox *et al.*, 2006; Cox *et al.*, 2008). Like the other alcohol-based fixatives, methacarn preserves tissues by coagulation and not by cross-linking proteins. This method does not mask antigenic sites, making the immunological studies less

complicated. This fixative also preserves the tissue morphology, and is not aggressive on tissue membrane (Shibutani *et al.*, 2000).

On the other hand, RNA*later* is a safe, non-toxic solution that is stored at room temperature. RNA isolated from RNA*later* preserved tissue (normal skin after surgery for cutaneous neoplasia, brain, thyroid, lung, heart, skeletal muscle stomach, small and large intestine, pancreas, spleen, lymph node, prostate, liver, and tongue) was found to be intact through visual examination using Ethidium Bromide-stained denaturing gels, Northern Blot analysis, and expression microarray analysis (Florell *et al.*, 2001). They also investigated the properties of samples treated with RNA*later* and demonstrated that it was possible to generate sections for histological and IHC analyses which has retained the properties of the source tissue. Srinivasan *et al.* (2002) also demonstrated the quality of tissue retrieved from RNA*later* for those studies. The results reported by Srinivasan *et al.* (2002) were similar to those found where other methods for fixation were utilised. However, most other methods do not yield material of acceptable quality for all purposes, but rather only for one or a few of the desired purposes.

In the present study, Methacarn was used to fix the tissues together with Bouin's fixation and O.C.T. (Optimal Cutting Temperature) compound, RNA*later* and liquid nitrogen preservation, and all of them were assessed for RNA and structure quality. The need for high quality RNA for subsequent analyses was the main reason for an exhaustive assessment of the quality of the RNA isolation methods.

5.1.2. mRNA quantification methodology

Currently, the method preferred for the quantification of mRNA levels is the Reverse Transcription quantitative Polymerase Chain Reaction (RT-qPCR), due to its simplicity, sensitivity, and specificity. The RT-qPCR method used here comprised a reverse transcription (RT) reaction, followed by a PCR carried out in the presence of SYBR Green, a fluorescent dye which binds to double-stranded DNA (dsDNA). As

the PCR proceeds, dsDNA accumulates and a fluorescent signal is generated which can be detected by a qPCR instrument.

There are certain factors which contribute to variation in the results of these processes. Firstly, there is an unavoidable variation in RNA from the sample, such as the amount of starting material. Secondly, once the RNA is isolated it becomes unstable. Thirdly, the RT is dependent on the abundance of the target template (Bustin *et al.*, 2005), on differences between tissues or cells in overall transcriptional activity, and on enzymatic efficiencies. For these reasons, it is necessary to assess the quality of the RNA, and take any precautions to reduce degradation to the minimum once the RNA is isolated (Thellin *et al.*, 1999).

In order to control these variables, which are not the result of the experimental variables under examination, it is necessary to standardise the samples by quantifying the starting material, i.e. number of cells or quantity of RNA. However, this does not take the RNA quality and enzymatic efficiencies into account.

An appropriate solution is the normalisation of mRNA levels between different samples for an accurate comparison of transcription levels (Bustin *et al.*, 2009). The common practice today is the use of reference or housekeeping genes as internal control genes. The perfect internal control genes should have a stable expression in the samples to be analysed and a constant transcription in all types of cells at any point, or in response to experimental treatment. These genes should be subject to the same analysis as the genes to be quantified. This process allows the normalisation of differences in the amount and quality of the starting material and differences in RNA preparation and cDNA synthesis, since the abundance of internal control transcripts is quantified in the same material as the transcripts of interest (Nygard *et al.*, 2007).

Even the most commonly used genes, such as *GAPDH* and *ACTB*, have been shown to exhibit variation in expression and cannot act as valid controls in certain cases (Thellin *et al.*, 1999; Nygard *et al.*, 2007). For example, some reference genes are involved in basic metabolic functions, but if the energy source for a cell or animal changes then different metabolic pathways may be up or down regulated. Thus, due

to the precision required for the interpretation of the resulting data from the RT-qPCR analysis, both the correct selection of high quality internal controls and the accurate quantification of these genes are crucial (Nygard *et al.*, 2007). In the absence of a universal control gene, the use of multiple internal control genes is recommended (Vandesompele *et al.*, 2002). Consequently, the selection of these genes needs to be done carefully, according to the tissues to be studied (Vandesompele *et al.*, 2002; Hellemans *et al.*, 2007).

A study by Nygard *et al.* (2007) identified an appropriate set of internal control genes for pig tissues. They performed RT-qPCRs for 17 different porcine tissues (liver, kidney, thymus, adipose, cortex cerebra, cerebellum, hippocampus, lymph nodules, muscle, heart, skin, pancreas, bone marrow, bladder, lung, stomach, and small intestine) using nine different genes (*ACTB, B2M1, GAPDH, HMBS, HPRT1, RPL4, SDHA, TBP1, YWHAZ*), which had been used in previous studies as internal control genes. The method used in the choice of appropriate internal controls is described by Vandesompele *et al.* (2002) and used in other RT-qPCR studies (Meller *et al.*, 2005; Kuijk *et al.*, 2007; Tramontana *et al.*, 2008).

Vandesompele *et al.* (2002) employed ten commonly used internal control genes, selecting genes with different functions, and determined their expression levels in different cell lines, as well as in normal human tissues. Vandesompele *et al.* (2002) developed a Visual Basic Application for Microsoft Excel (GeNorm) which automatically calculates the gene-stability measure (*M*) for all internal control genes in a given set of samples. *M* was calculated as the average pair-wise variation of a particular gene with all other control genes. Genes with the lowest *M* values had the most stable expression. With this method, the best-performing internal control genes were found and their expression levels were used to calculate a normalisation factor. Vandesompele *et al.* (2002) recommended a minimum of three internal control genes to calculate the normalisation factor, but allowed some flexibility about the number, depending on the sample variation, the genes under study, and the amount of RNA available. They also expressed the opinion that it was unnecessary to use a great number of internal control genes where the target genes are few or where all control genes are relatively stable expressed, without changing the normalisation factor

through the inclusion of these genes. The need to investigate every individual case without making any assumptions was stated.

5.2. Materials and Methods

Initially, reproductive tissues (placenta, endometrium, and whole utero-placental units) were collected from one sow (LW x LR crossbreed) (id 574) at day 42 of pregnancy, and from a second sow at day 41 (id 509). Seven further pigs were slaughtered later. Most of the tissues from the first two pigs and some of the others (some extra tissue was also collected) were used in an initial validation and optimisation study, thus, the quality of the material, and the optimal techniques for tissue preservation, RNA extraction and RT-qPCR techniques were verified. The analysis of the two first pigs allowed the refinement of techniques to be used in the collection of the seven remaining pigs and the robustness of the techniques to be tested. The key facts relevant to the analyses described in this Chapter, concerning the animals sampled – parity, stage of pregnancy and LS – are shown in Table 5.1. Further details are reported in Appendix 4.

Pig id	**574**	**509**	**Y24**	**W12**	**Y22**	**W2**	**W8**	**Y26**	**W7**
Pig number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Parity number	1	1	2	1	1	2	2	2	2
Stage of pregnancy (days)	42	41	45	46	44	44	43	41	42
Litter size	15	12	16	16	13	18	17	10	12

Table 5.1 Status of animals sampled. This table summarises information for sows/gilts from which tissues were collected, including animal id, designated pig number, parity number, stage of pregnancy in days, and litter size.

5.2.1. Tissue collection from Large White x Landrace crossbred gilts/sows

LW x LR crossbred pregnant gilts/sows were sacrificed at The Roslin Institute's Large Animal Unit and reproductive tissues collected and preserved for use in

multiple experiments. For this purpose tissues were protected from RNases at all times by wiping the instruments and trays before use with RNase*Zap* (Applied Biosystems). The collection of tissues, in this experimental animal facility, was done as quickly as possible after the death of the animal, to reduce RNA degradation (a maximum of 1 hour).

The gilts/sows were sedated by intra muscular injection of a mixture of Ketamine (7.25 mg per Kg) (W and J. Dunlop Ltd., Dumfries, UK) and Azaperone (1.25 mg per Kg) (W and J. Dunlop Ltd.). Following sedation (12 to 15 minutes to ensure maximum effect) the pigs were euthanised with sodium pentobarbitone 20% w/v (W and J. Dunlop Ltd.) at a dose rate of 0.4 ml/kg delivered by an intravenous injection through a cannula placed in an ear vein. After death was confirmed, the pig was moved to a table where the abdomen was washed (e.g. with Savlon (W and J. Dunlop) diluted 1:14 with water) and a mid-ventral incision through the skin, fat, and body wall was made. Using gloved hands rinsed with RNase*Zap*, the pregnant reproductive tract was lifted out of the body cavity and removed by cutting through the vagina, ensuring that both ovaries were retained. The tract was collected in a dissecting tray and transferred to the dissecting area.

The position of the smallest foetus was identified by palpating the tract, and this unit was dissected first. Subsequently, a normal sized foetus was identified and dissected. For each feto-placental unit studied, a shallow incision through the uterine wall and endometrium was made. These tissues were peeled open to expose and rupture the allantoic sac. The amnion was opened, the umbilicus cut close to the foetus and the foetus removed and weighed. Sections of 'whole uterus' (chorioallantois, endometrium, and myometrium), defined also as an utero-placental unit, endometrium, and chorioallantois (defined hereafter as placenta) were collected from the area that was in contact with the amnion *in situ* (Figure 5.1).

After the collection of tissues from both a small and normal foetus, all remaining feto-placental units were dissected and foetus weighted starting from Left 1, where feto-placental unit 1 is closest to the ovary, to get the correct position of each foetus,

without harvesting any tissue from them, and then Right 1. All available information about the gilt/sow, and foetuses was recorded on a spreadsheet.

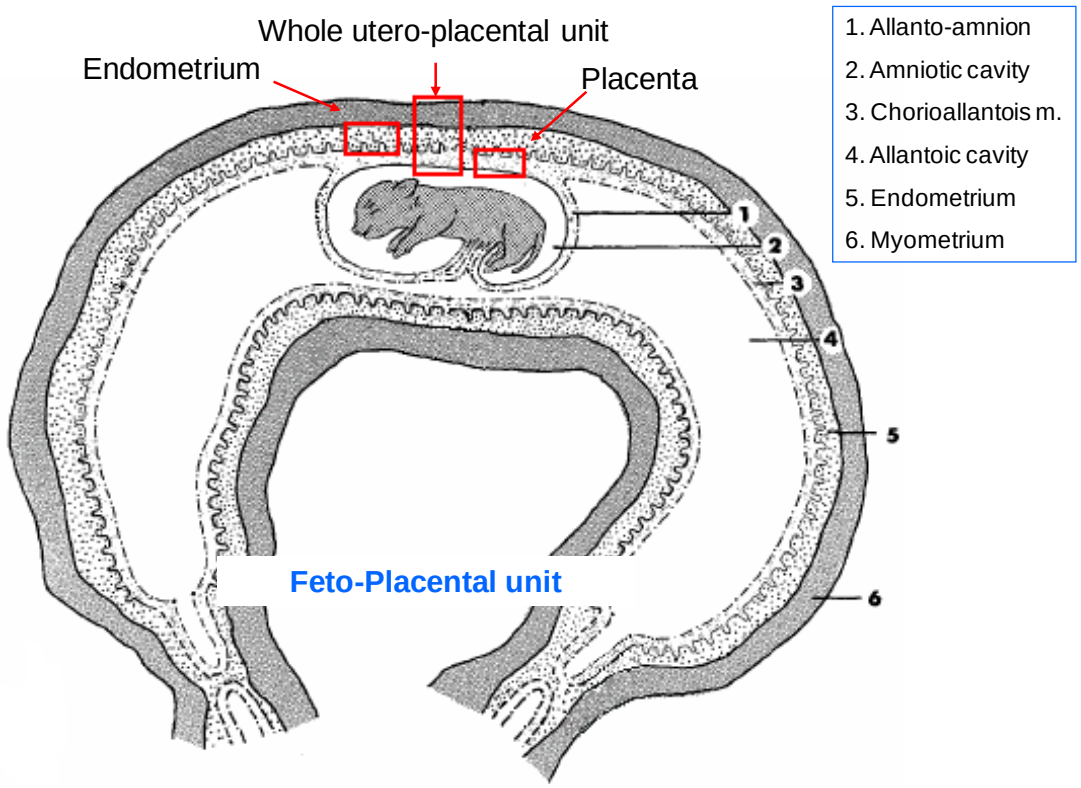

Figure 5.1 Diagram representing a feto-placental unit where the tissues collected are indicated and defined. Adapted from Rice *et al.* (1991).

5.2.2. Tissue preservation and processing

Each tissue of interest was removed from the animal and trimmed into appropriately sized pieces, depending on the preservation method. A piece of each tissue was cut into small pieces and placed in a cryovial (Camlab, Cambridge, UK) and this was inserted into a self seal bag with holes and the appropriate paper label. The bag and its contents were snap frozen in liquid nitrogen (BOC, Edinburgh, UK) and transferred to dry ice when thoroughly frozen. The samples were stored at -80ºC.

A small piece of each tissue was cut and placed in a Lysing matrix D tube (MP Biomedicals, Illkirch, France) with 1 ml of RNA-Bee RNA isolation Reagent (AMS Biotechnology, Abingdon, UK) and was homogenised using a Fast Prep (BIO101 –

FP120, Thermo electron corporation, Ohio, USA) at speed 4.0 to 5.0 for 20 seconds. The tubes were placed in dry ice in order to freeze them and then stored at -80ºC until processed.

Another piece of tissue was placed on a small drop of O.C.T. compound (VWR international, BDH Prolabo, Leicestershire, UK) on a coin, then covered in more O.C.T. compound and placed on a rack just above the surface of liquid nitrogen. Once frozen it was transferred to a bag and further frozen by immersing in liquid nitrogen. The samples were subsequently stored at -80ºC.

A sample of each tissue of interest was trimmed to ≤ 0.5 cm in any single dimension and placed in a plastic universal (Sterilin, Newport, UK – part of Thermo Fisher) containing 5-10 volumes of RNA*later*® Solution (Applied Biosystems/Life technologies, Paisley, UK), an aqueous solution that is designed to stabilise and protect RNA in fresh tissue samples, and stored at 4ºC overnight. After incubation at 4ºC overnight the RNA*later* samples were stored at -20ºC or at -80ºC after removing the RNA*later*. RNA*later* was not initially considered as an option in this study, due to the previous efficient use of liquid nitrogen as a preservation method for RNA. Taking into account the results obtained from initial collections (pig 1 and 2) and the efficiency demonstrated in other studies, RNA*later* was incorporated into the study for subsequent collections.

A representative square of tissue, or rectangular section for 'whole uterus', was cut for embedding and placed in a plastic universal containing approximately 15 ml of Methacarn: 60% (v/v) Methanol (Fisher Scientific Ltd.), 30% (v/v) Chloroform (Fisher Scientific Ltd.) and 10% (v/v) Acetic Acid (VWR international). The samples were left at 4ºC for 24 hours/overnight before processing. After this, the Methacarn was removed and the rest of the fixation procedure was performed. In the cases when it was not possible to perform the next step after the Methacarn fixation (maximum 24 hours), tissues were transferred to 100% ethanol (Fisher Scientific Ltd.) for no more than 24 hours.

Another sample was placed in a plastic universal containing approximately 15 ml of Bouin's solution (Sigma-Aldrich). The container was left at room temperature for 24 hours. After incubation, the Bouin's solution was removed from the fixed samples, and the samples were transferred to 70% ethanol and stored at room temperature until needed.

After fixation in Methacarn, Bouin's, 100% ethanol or 70% ethanol, tissue processing/embedding cassettes (Histosette®, Simport, Beloeil, Canada) were labelled for each tissue piece and the corresponding tissues placed in them. The tissues in the cassettes were washed three times for 1 hour in 100% ethanol at 4ºC, for 1 hour in Xylene (Fisher Scientific Ltd.) at room temperature and three times for 30 minutes in Xylene at room temperature. Finally, they were washed three times for 1 hour in paraffin wax (VWR international) at 60ºC. After the washes, the cassettes were opened and the samples were placed in a mould. The cassette was placed on the mould over the tissue and filled with paraffin. They were left to cool down for at least 2 hours or overnight.

5.2.3. Spectrophotometry

The concentrations of the isolated RNA samples were measured with a Nanodrop ND-1000 (Labtech International Ltd.), a UV/Vis full-spectrum (220-750) spectrophotometer. Prior to measurement of the sample, the spectrophotometer was calibrated with RNase free water or elution solution. The concentration of the sample (1.2 µl) was measured in ng/µl. The OD (Optical density) 260/280 nm and OD 260/230 nm ratios were also calculated and a graph of absorbance over the wavelength range 220-350 nm was displayed for each measurement. The ratio of sample absorbance at 260 and 280 nm was used to assess the purity of RNA. A ratio of ~2.0 is generally accepted as "pure" for RNA. If the ratio is appreciably lower, it may indicate the presence of protein, phenol or other contaminants, which have high absorption at or near 280 nm. The 260/230 ratio is a secondary measure of nucleic acid purity, and they are often higher than the respective 260/280 ratio values. They are commonly in the range of 1.8-2.2. If the ratio is appreciably lower, this may

indicate the presence of co-purified contaminants. A blank solution was also measured as a control.

The RNA concentrations, as estimated with the Nanodrop, were used to ensure, to the extent possible, that equal amounts of each sample were used in further steps. The RNA concentrations varied greatly among samples due to the different amount of tissue used for each isolation. The OD 260/280 nm and OD 260/230 nm ratios were considered adequate around ~ 2.0 and in the range of 1.8-2.2, respectively. If a sample did not fulfil this criterion (Table 5.2), it was cleaned up with a column (5.2.7), or the RNA isolation was repeated.

Evaluation	OD 260/280	OD 260/230	Quality
✓✓	~ 2.0	1.8-2.2	High
✓	~ 2.0	1.6-1.79 and 2.21-2.4	Acceptable
✓✗	~ 2.0	0-1.59 and 2.4-more	Low
✗	other	other	None

Table 5.2 Definition of threshold for evaluation of spectrophotometry results. Indication of the symbols to define validation of the results.

5.2.4. RNA quality control by electrophoresis

The quality of the isolated RNA was assessed with a RNA 6000 LabChip kit (Applied Biosystems) on an Agilent 2100 Bioanalyser (Agilent Technologies, Edinburgh, UK) (Service offered by ARK-Genomics, The Roslin Institute). This system allows the integrity and purity of RNA samples to be assessed using fluorescence detection, monitoring the fluorescence between 670 and 700 nm for RNA. The electrodes of the bioanalyser were decontaminated and the chip was prepared before loading 1 µl of each sample together with 1 µl of a RNA 6000 Nano marker (Applied Biosystems). This kit allowed the quantification and integrity analysis of total RNA measurement by an RNA integrity number or RIN, as well as the visualisation of rRNA by a gel-like image which showed the two ribosomal bands for 18S and 28S, indicating their intact presence or not. If these two rRNA are found intact, the integrity of the rest of the RNA is inferred. The RIN ranges from 1

to 10, with 10 indicating the highest quality. For this study the RIN minimum was set at 7, if a sample failed to reach this value (Table 5.3), the isolation was repeated.

Evaluation	RIN	Quality
×	0 - 4.9	None
✓	5 - 6.9	Acceptable
✓✓	7-10	High

Table 5.3 Evaluation criteria for RNA integrity.

5.2.5. RNA isolation from tissues embedded in paraffin

RNA was isolated from tissues embedded in paraffin (Bouin's and Methacarn) with the RecoverAll™ Total Nucleic Acid Isolation, optimised for FFPE (Formalin-fixed, Paraffin Embedded) Samples kit (Applied Biosystems). This kit included digestion buffer, protease, isolation additive, filter cartridge, collection tube, Wash 1, Wash 2/3, 10x DNase buffer, DNase, elution solution, and nuclease-free water. Four 20 µm sections were taken from each sample with a microtome HM 325 (MICROM international GmbH, Walldorf, Germany) and put into a 1.5 ml autoclaved microcentrifuge tube (Sarstedt Ltd., Leicester. UK). One ml of Xylene was added to each sample and mixed briefly by vortexing. The mixture was heated for 3 minutes at 50ºC to melt the paraffin, and centrifuged at room temperature for 2 minutes at maximum speed. The Xylene was removed, and the pellet was washed with 1 ml of 100% ethanol at room temperature, mixed, and centrifuged for 2 minutes at room temperature at maximum speed to remove any Xylene. The ethanol was discarded and the wash was repeated for 2 minutes. The ethanol was again removed, and the tube was centrifuged briefly to remove any remaining ethanol. The pellet was air dried for 10-15 minutes. The protease digestion was performed by adding 400 µl of digestion buffer plus 4 µl of protease to each sample. The tube was swirled gently to mix and briefly centrifuged. The sample was incubated for 3 hours at 50ºC in a water bath.

For the nucleic acid isolation, 480 µl of isolation additive was added to each sample, and mixed by vortexing, resulting in a white solution. One thousand one hundred µl

of 100% ethanol was added and each sample was mixed by pipetting carefully, resulting in a clear solution. Seven hundred µl of the sample/ethanol mixture was pipetted onto a filter cartridge placed in a collection tube. The mixture was passed through the filter by centrifuging at 10,000 x *g* for 30-60 seconds. The flow-through was discarded and the rest of the mixture was pipetted on to the filter and passed through as before. Wash 1 (700 µl) was added to the filter, the tube was centrifuged for 30 seconds at 10,000 x *g* and the flow-through was discarded. Wash 2/3 (500 µl) was added to the filter, and centrifuged for 30 seconds to remove residual fluid.

In order to remove any contaminating DNA, a DNase master mix containing 6 µl 10x DNase buffer, 4 µl DNase, and 50 µl nuclease-free water was prepared and added to the centre of the filter, the tube was capped and incubated for 30 minutes at room temperature. Wash 1 (700 µl) was added to the filter, and the tube was incubated 30-60 seconds at room temperature. The filter and collection tube were centrifuged for 30 seconds at 10,000 x *g*. The flow through was discarded and 500 µl of wash 2/3 was added to the filter, and centrifuged for 30 seconds at 10,000 x *g*. The flow through was discarded and this last step was repeated. Any remaining residual fluid left was removed by centrifuging for 1 minute at 10,000 x *g*.

The filter was transferred to a fresh collection tube and 30 µl of elution solution pre-heated to 95ºC, was added to the centre of the filter for the final nucleic acid purification. This was allowed to sit at room temperature for 1 minute and then centrifuged for 1 minute at maximum speed. The elution step was repeated with another 30 µl of elution solution. The final result was 60 µl of collected eluate which contains the RNA. The concentration of RNA in the collected sample was analysed by spectrophotometry and the sample was stored at -80ºC.

5.2.6. RNA isolation from (Frozen) tissues (Method 1)

Approximately 50 mg of tissue was trimmed and homogenised in Lysing matrix D tubes (MP Biomedicals) with 1 ml of RNA-Bee RNA isolation Reagent (AMS Biotechnology) using a Fast Prep lysis instrument at speed 4.0 to 5.0 for 20 seconds.

The homogenisation was repeated once more only if little or none of the tissue had been homogenised with this first pulse. The tissues were kept on dry ice until just before homogenisation. The homogenate was removed from the Lysing matrix D tubes and put into a clean autoclaved tube (Sarstedt Ltd.), 0.2 ml of chloroform was added, and the sample was shaken vigorously for 15-30 seconds. The sample was incubated on ice for 5 minutes. After the incubation, the sample was centrifuged at 12,000 x *g* for 15 minutes at 4ºC.

In this last step two phases were formed. The upper one, an aqueous phase of around 500 µl which contained the RNA, was transferred to a fresh tube. The lower phase was discarded. A volume of Isopropanol (Fisher Scientific Ltd.), equivalent to the volume of the aqueous phase transferred, was added, and the sample was held at room temperature for 5-10 minutes to precipitate the RNA. After this period of incubation the tube was centrifuged at 12,000 x *g* for 10 minutes at room temperature. The resulting pellet was washed once with 1 ml of 75% ethanol, mixed and centrifuged at 7,500 x *g* for 5 minutes at room temperature. The pellet was air dried at room temperature for 5-10 minutes, and dissolved in an appropriate volume of water, depending on the size of the pellet. The RNA was quantified, and its quality assessed by spectrophotometry and electrophoresis, respectively, and stored at -80ºC. If the results from the spectrophotometry were not adequate, the samples were cleaned using an RNeasy Mini kit (5.2.7). Equally, if the results from the electrophoresis were not adequate the isolation was repeated.

5.2.7. Clean up column

A Qiagen RNeasy Mini kit (QIAGEN, West Sussex, UK) clean up column was used to clean up the isolated RNA. This kit contained buffer RPE, buffer RW1, buffer RLT, and RNase-free water. Each RNA sample was made up to 100 µl with RNase-free water, 350 µl of Buffer RLT was added and the sample was mixed thoroughly. Absolute ethanol (250 µl) was added and the sample mixed thoroughly by pipetting. The sample was loaded onto a column in 700 µl aliquots, where the RNA was captured, the tube was centrifuged for 15 seconds at ≥8,000 x *g*, and the flow-

through discarded after each loading. The column was transferred to a fresh collection tube, washed with 500 µl Buffer RPE, and centrifuged at ≥8,000 x *g* for 15 seconds. The flow through was discarded and the column was placed in the same collection tube. The column was washed with 500 µl of Buffer RPE and centrifuged for 2 minutes at ≥8,000 x *g*. The column was then transferred to a new 1.5 ml collection tube and centrifuged at full speed (16,100 x *g*) for 1 minute. The column was transferred to a fresh collection tube and 2 x 50 µl of RNase-free water was added onto the column membrane. The tube was centrifuged for 1 minute at ≥8,000 x *g* each time. Both elutes were mixed, and then the RNA concentration was measured on the Nanodrop and the samples were stored at -80ºC.

5.2.8. RNA isolation from tissues potentially containing high amounts of lipophilic components (Method 2)

Approximately 50 mg of tissue were homogenised in Lysing matrix D tubes with 1 ml of RNA-Bee RNA isolation Reagent in a FastPrep (2 x 20 seconds at speed 6.0), and placed on ice. The homogenate was transferred to an autoclaved tube and centrifuged at 4ºC for 10 minutes at 12,000 x *g*. The supernatant was transferred to a fresh tube and kept at room temperature for 5 minutes. Chloroform (200 µl) was added, and the tube was shaken for 15 seconds and kept at room temperature for 10 minutes. The sample was centrifuged at 12,000 x *g* at 4ºC for 15 minutes. The upper phase, approximately 500 µl, was transferred to a fresh tube, one volume of 85% ethanol was added and the tubes were vortexed. The samples were loaded onto an RNeasy column in 700 µl aliquots. The columns were centrifuged at 8,000 x *g* for 15 seconds and the flow through discarded after each loading. The column was washed and centrifuged at 8,000 x *g* with 350 µl of Buffer RW1 for 15 minutes, with 500 µl Buffer RPE for 15 seconds, and with 500 µl Buffer RPE for 2 minutes. The column was transferred to a new tube and centrifuged at full speed for 1 minute. The column was transferred to a new tube and the RNA was eluted in 2 x 50 µl of RNase-free water (the water was added, the column stood for 1 minute, then centrifuged for 1 minute at 8,000 x *g*). This method had been used previously with pig placental tissue (McNeil *et al.*, 2007)

5.2.9. Final RNA isolation method (Method 3)

The tissues preserved in liquid nitrogen or RNA*later* were cut and a piece weighing ~30 mg (50 mg maximum) was placed in an RNase-free eppendorf tube (Thistle Scientific, Glasgow, UK) and frozen in dry ice to improve the homogenisation. The tissue was homogenised in RNA-Bee in Lysing matrix D tubes using a Fast Prep (20 seconds at speed 4.0). The homogenate, including some tissue pieces, if present, was transferred (~1 ml) to a fresh RNase-free tube. The sample was centrifuged at 12,000 x *g* for 10 minutes at 4ºC. The supernatant (~1 ml) was transferred to a fresh tube and held at room temperature for 5 minutes. Chloroform (200 µl) was added, the tube was shaken vigorously for 15-30 seconds, and then kept on ice for 10 minutes. The sample was centrifuged at 12,000 x *g* for 15 minutes at 4ºC. About 500 µl or less was removed from the upper, aqueous phase to a fresh tube. One volume (aqueous phase) of 70% ethanol was added and mixed by pipetting.

Up to 700 µl of the sample were transferred to an RNeasy column placed in a 2 ml collection tube and centrifuged for 15 seconds at ≥8,000 x *g*. The flow-through was discarded and the column was placed in the same collection tube. The remains of the sample were transferred, and the centrifugation was repeated. The flow-through was discarded and the column was placed in the same collection tube. Buffer RW1 (700 µl) was added to the column and it was centrifuged for 15 seconds at 8,000 x *g*. The flow-through was discarded and the column was placed in the same collection tube. Buffer RPE (500 µl) was added to the tube and it was centrifuged for 15 seconds at ≥8,000 x *g*. The flow-through was discarded and the column was placed in the same collection tube. Buffer RPE (500 µl) was added and the tube was centrifuged for 2 minutes at ≥8,000 x *g*. The column was transferred to a new 2 ml collection tube and centrifuged at maximum speed (16,100 x *g*) for one minute to purge the column of Buffer RPE. The column was transferred to a new 1.5 ml RNase-free collection tube and 30-50 µl of RNase-free water was pipetted onto the column membrane. The tube was centrifuged for 1 minute at ≥8,000 x *g* to elute RNA. The elution process was repeated (in the same tube) with a further 30-50 µl of RNase-free water. At the end of the elution process the tube was vortexed to mix

both eluates. The RNA was stored at -80ºC. This RNA was quantified by spectrophotometry and its quality checked using the Agilent Bioanalyser.

5.2.10. RNA isolation method with TRIzol (Method 4)

This method was the same as Method 3 (5.2.9) except for swapping RNA-Bee for TRIzol® Reagent (Invitrogen, Parsley, UK).

5.2.11. RNA isolation method with Ultraspec (Method 5)

The first steps of this method were the same as Method 3 (5.2.9) but using Ultraspec II (Biotecx, Houston, USA) instead of RNA-Bee, up to the collection of the upper aqueous phase after the centrifugation for 15 minutes at 4ºC. Once this phase was in a fresh tube, half a volume of Isopropanol was added and the tube was mixed. A twentieth volume (aqueous phase) of resin was added and mixed. The tube was centrifuged for 1 minute at maximum speed (16,100 x *g*). One ml of 75% ethanol was added, mixed by vortexing for 30 seconds, and centrifuged for 30 seconds at maximum speed. The supernatant was removed and the ethanol wash was repeated. The supernatant was removed and the tube was briefly re-centrifuged to remove any traces of ethanol. The pellet was air dried for approximately 5 minutes and a twentieth volume of water was added, the tube was vortexed and centrifuged for 1 minute at maximum speed. The supernatant was transferred to a new tube. Then, the concentration was measured by spectrophotometry and the samples were stored at -80 ºC.

5.2.12. Reverse Transcription reactions (using SuperScript III)

Random primers were annealed to a template RNA and a reverse transcriptase enzyme was used to synthesise cDNA. The RNA sample volume required for a specific weight of RNA was calculated with the results from the spectrophotometry. The same weight of RNA was required for each sample in the group, in order to be able to compare gene expression levels.

The reaction was performed in a 200 µl PCR tube (Axygen, Union City, CA, USA-supplied by Thistle Scientific) and consisted of the RNA sample (1.25 µg), 1 µl of random primers (250 ng/µl; Promega Corporation, Southhampton, UK), 1 µl of dNTPs (10 mM; Invitrogen), and RNase-free water up to a volume of 13 µl. The tubes were mixed, centrifuged briefly and incubated for 5 minutes at 65 ºC, then cooled to 4ºC for at least 2 minutes in a Thermocycler (Tprofessional, Biometra, Goettingen, Germany).

After this first incubation, 4 µl of 1st strand buffer (Invitrogen), 1 µl of 0.1 M DTT (Invitrogen), 1 µl of RNase inhibitor (40 units/µl; RNasin, Promega Corporation), and 1 µl of Superscript III (200 units/µl; Invitrogen) were added to the 13 µl mix as a master mix. To check for genomic DNA contamination, one sample containing no Superscript III was included per group (in this case per pig). The tubes were centrifuged briefly and heated to 25ºC for 5 minutes, 50ºC for 60 minutes, and 70ºC for 15 minutes in the thermocycler. The samples were stored at -20ºC until required.

5.2.13. PCR

The samples of cDNA were tested for two different genes; the one of interest, *SPP1* where the primers were as described by Hellemans *et al.* (2007) forward TTG GAC AGC CAA GAG AAG GAC AGT and reverse GCT CAT TGC TCC CAT CAT AGG TCT TG (GenBank accession number X16575) (Sigma-Aldrich), and expected to yield a 120 bps PCR fragment; and *ACTB* where the primers were forward GAG AAG CTG TGC TAC GTC GC and reverse CCA GAC AGC ACT GTG TTG GC, (GenBank accession number DQ452569.1) as an internal control gene and expected to yield a 259 bps fragment.

A master mix was prepared to provide 2.5 µl of 10x reaction buffer (Roche, included with the taq kit), 0.5 µl of 10 mM dNTPs (Invitrogen), 2 µl of each primer (5 µM), 1 µl of Taq/TaqStart mixture, and 16 µl of nuclease-free water, per tube. The Taq/Taq Start mixture consisted of one part of Taq DNA polymerase (Roche), one part of TaqStart antibody (Clontech, Saint-Germain-en-Laye, France), mixed with 3

parts of TaqStart dilution buffer (Clontech) incubated at room temperature for 5 minutes. The use of this Taq/TaqStart mix is expected to help minimise the formation of primer-dimers. Apart from the tissue samples and the RT blank (cDNA control that contains RNA, but no Superscript enzyme), a negative control was included, containing the master mix but no cDNA.

Twenty four µl of the master mix were added to each tube, together with 1 µl of cDNA for each sample and 1 µl of nuclease-free water to the negative control. The tubes were centrifuged briefly and placed in the thermocycler (Tprofessional, Biometra). The following program was run: one step at 95ºC for 4 minutes 30 seconds, a second cycle consisting of one step at 95ºC (denature) for 30 seconds, a second step at 60ºC (annealing temperature (Tm), specific for each primer) for 30 seconds, and a last step at 72ºC (extension) for 30 seconds, repeated 29 times more, amplifying the region between the primers, and a final step at 72ºC for 4 minutes and 30 seconds. In order to check the results, 5 µl of the PCR product were run on an agarose gel and visualised.

5.2.14. Agarose gel

The DNA samples were visualised using a 3% (w/v) agarose gel (Sigma-Aldrich) in 1x TAE buffer prepared from 50x TAE (Tris Base (484 g), Glacial Acetic acid (114.5 ml), and 5 M EDTA pH 8.0 (200 ml), dissolved in 1 litre of dH_2O and made up to 2 litres with water once dissolved). In a conical flask, the agarose was melted in the buffer and 10 µl of DNA stain SYBR® Safe (Invitrogen) was added per 100 ml of gel. The gel was allowed to set in the tray, then it was placed in the tank and covered with 1x running buffer. Five µl of the samples from the PCR, mixed with 1 µl of the loading buffer (Promega blue/orange) for a 1x buffer concentration, were loaded in a well of the gel and a current between 100 and 150 volts with a Power Pac (BIO-RAD 300, Bio-Rad Laboratories Ltd.) was applied. A set of size markers (Promega Corporation) was included. The DNA migrated towards the positive electrode at a rate proportional to its size. SYBR® Safe DNA gel stain (Invitrogen)

stained the DNA allowing it to be visualised and photographed under U.V. or blue (Safe Imager™, Invitrogen) light.

5.2.15. Quantitative PCR

The cDNA samples prepared previously from the RNA isolated from the liquid nitrogen preserved samples for pig 1 and 2 were diluted 1:8 with nuclease-free water in the appropriate proportion. A pool of cDNA was prepared by taking a small volume from each undiluted sample as standards. An equal volume of nuclease-free water was added to this pool (1:2). The pool was mixed and double diluted to give a 1:4, 1:8, 1:16, 1:32, and 1:64 series of dilutions of the original, to obtain the standard curve.

A master mix was prepared to give the following volumes per tube/well: 12.5 µl of Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen), 0.1 µl of ROX reference dye, 2 µl of Forward primer (5 µM), 2 µl of Reverse primer (5 µM), and 3.4 µl of Nuclease-free water. Twenty µl of master mix was added to each tube/well, in a 96 well plate (ABgene®, Thermo Scientific). A blank (nuclease-free water) or NTC (no template control) was added in the first well and 5 µl of each of the standards was added to the appropriate following wells. In a second step, 5 µl of each diluted cDNA were added to the appropriate well, followed by a RT blank for each group of samples (no RT). The samples, standards and blanks were measured in duplicate. The strategy used was the sample maximisation technique, where as many samples as possible were loaded in the same plate and different genes were in different runs. This allowed the comparison of expression level in different samples for the same gene and minimised the technical variation from run to run between the samples (Nygard *et al.*, 2007).

The plate was capped with optical caps (Applied Biosystems), and briefly centrifuged to bring everything to the bottom of the wells and to remove bubbles. The plate was placed in a Stratagene MX3000P qPCR instrument (Agilent Technologies) and the settings were adjusted for analysis with the Invitrogen

SYBRGreen thermal profile (Figure 5.2). A first cycle consisting of 2 minutes at 50ºC and 2 minutes at 95ºC was set. This cycle was followed by 40 cycles of 15 seconds at 95ºC and 30 seconds at 60ºC. The final cycle consisted of one minute at 95ºC, 30 seconds at 60ºC, 15 seconds at 95ºC and 30 seconds at 25ºC.

The results were saved in a specific format, compatible with the Stratagene MX3000 instrument, thus the Stratagene MXpro software (Agilent technologies) was required to visualise the results. Once the file was opened the wells were labelled as standards, unknowns for the samples, NTC (no template control) or no RT blanks. For the standards, a value relative to the sample dilution was assigned (4, 2, 1, 0.5, 0.25, and 0.125) in order to create a standard curve. The results were visualised in different ways. The amplification plot was used to visualise each sample and standard duplicates in order to check for possible differences between duplicates or any samples with errors. The amplification plots indicated where the samples crossed the threshold or cycle threshold (Ct). The standard curve, created with the standards, indicated the efficiency of the qPCR that should be approximately 90-110%. For a good standard curve, the duplicates of the standards should be similar. The standard curve was visualised in a plot where the standards defined the curve and the samples were represented by dots. This was verified in order to make sure the samples were contained within the range of the curve, indicating the adequacy of the standards. The text report option gave the results in a table with the labels for each well, the values for the standards, and the Ct result for each sample, which was exported in Excel format. The Ct indicated the cycle where the fluorescence level for that sample crossed the threshold.

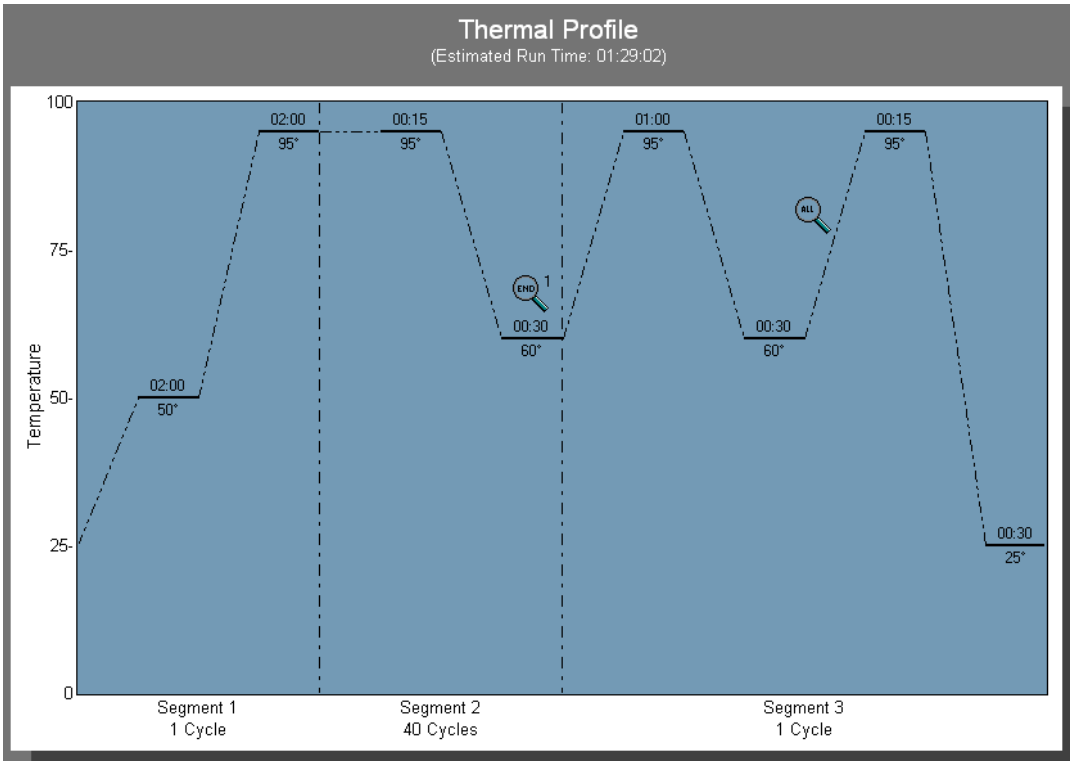

Figure 5.2 Invitrogen SYBRGreen Thermal profile. This figure illustrates the amplification process performed in the qPCR machine. It shows the different cycles performed, and the corresponding temperature and time of each cycle. Briefly, first 2 minutes at 50ºC, 2 minutes at 95ºC, 40 cycles of 15 seconds at 95ºC, and 30 seconds at 60ºC. The final cycle consisted of one minute at 95ºC, 30 seconds at 60ºC, 15 seconds at 95ºC and 30 seconds at 25ºC.

5.2.16. Sectioning

Sections were taken from the tissues embedded in paraffin with a Microtome (HM 325; MICROM international) using disposable knives, Shandon MB-35 (Thermo Scientific). Sections were cut at 5 µm and lifted over a water bath at 48ºC. Polysine slides (Thermo Scientific) were used to lift out the tissue from the water bath, introducing them at a 45º angle. The slides were left to dry for 10 minutes and then placed in the oven at 60ºC for at least 2 hours or overnight, in order to allow the polysine to denature and the sections of tissue to adhere to the slide.

5.2.17. Haematoxylin and Eosin Staining of Paraffin Embedded and OCT Sections

The first step for the Haematoxylin (VWR International) and Eosin (Sigma-Aldrich) staining, consisted of a number of washes for which the sections were placed in a staining rack. The paraffin wax was removed by treating it with Xylene for 5 minutes and the sections were rehydrated before staining by treating them with 100% ethanol twice for 5 minutes, with 90% ethanol for 5 minutes, and with 70% ethanol for 5 minutes. After re-hydration the sections were washed in dH_2O for 5 minutes. The next step was the Haematoxylin staining for 5 minutes. This was also the first step for the O.C.T. sections. The sections were moved to running cold tap water for 5 minutes to remove the haematoxylin excess. Sections were treated for 2 minutes in Eosin. They were dipped 10 times in dH_2O, 70% Ethanol, 90% Ethanol, and 100% Ethanol. To finish, the sections were treated with Xylene for at least 10 minutes.

The slides were drained onto paper tissue and a drop of DPX mountant (Sigma-Aldrich) was added to the section under the fume hood and a cover slip was placed carefully on top, avoiding the formation of any bubbles. The slides were allowed to dry before being viewed under an appropriate microscope (Nikon Optiphot2, Surrey, UK). Pictures were taken with a Scion camera (Scion Corporation, Maryland, UK) in order to analyse the structures and the difference between the treatments. This staining allowed the structure of the section to be seen clearly. Haematoxylin binds to basophilic structures, most obviously the chromatin in the nucleus (blue/purple). Eosin binds to eosinophilic structures, mostly cytoplasmic (pink/red) and extracellular proteins (deep pink).

5.2.18. Selection of internal control genes for the qPCR

There was a need to identify adequate internal control genes for the tissues of interest. In order to choose the proper internal controls for the RT-qPCR analysis, the specific primers for nine different genes in pig (Invitrogen) (Tables 5.4 and 5.5) (Vandesompele *et al.*, 2002) were tested with six samples from the same pig. The

compatibility and specificity of these primers were tested on samples from the target tissues and to determine the optimal conditions of these primers, PCRs were run for the nine genes listed in Table 5.4, plus *SPP1*. The cDNA for these samples was prepared as indicated in 5.2.12. The annealing temperature (Tm) for these primers was initially set at 60ºC and for 30 cycles, in a similar process to that described in 5.2.13.

Gene	Forward	Amplicon Size	Tm (°C)
ACTB	CACGCCATCCTGCGTCTGGA	379	63
	AGCACCGTGTTGGCGTAGAG		
B2M1	CAAGATAGTTAAGTGGGATCGAGAC	161	58
	TGGTAACATCAATACGATTTCTGA		
GAPDH	ACACTCACTCTTCTACCTTTG	90	60
	CAAATTCATTGTCGTACCAG		
HMBS	AGGATGGGCAACTCTACCTG	83	58
	GATGGTGGCCTGCATAGTCT		
HPRT1	GGACTTGAATCATGTTTGTG	91	60
	CAGATGTTTCCAAACTCAAC		
RPL4	CAAGAGTAACTACAACCTTC	122	60
	GAACTCTACGATGAATCTTC		
SDHA	CTACAAGGGGCAGGTTCTGA	141	58
	AAGACAACGAGGTCCAGGAG		
TBP1	AACAGTTCAGTAGTTATGAGCCAGA	153	60
	AGATGTTCTCAAACGCTTCG		
YWHAZ	TGATGATAAGAAAGGGATTGTGG	203	60
	GTTCAGCAATGGCTTCATCA		

Table 5.4 Primers of the internal controls genes. This table shows the names of the genes, the forward and reverse primer for each gene, the expected amplicon size in bps, and the annealing temperature (Tm) for each pair of primers (Nygard *et al.*, 2007).

Gene	Name	Accession Number	Function
ACTB	β-Actin	DQ845171	Involved in cell motility, structure and integrity
B2M1	β- 2- microglobulin	DQ845172	Cytoskeletal protein involved in cell locomotion
GAPDH	Glyceraldehyde-3- phosphate dehydrogenase	DQ845173	Carbohydrate metabolism
HMBS	Hydroxymethylbilane synthase	DQ845174	Heme biosynthesis
HPRT1	Hypoxanthine phosphoribosyltransferase I	DQ845175	Purine ribonucleoside salvage
RPL4	Ribosomal protein L4	DQ845176	Structural constituent of ribosome
SDHA	Succinate dehydrogenase complex subunit A	DQ845177	Tricarboxylic acid cycle
TBP1	TATA box binding protein	DQ845178	Transcription initiation from RNA polymerase II promotor
YWHAZ	Tyrosine 3-monooxygenase/tryptophan 5-monoxigenase activation protein zeta polypeptide	DQ845179	Protein domain specific binding

Table 5.5 Information on the internal control genes including the full name of the gene, the accession number in Genbank and the function of the gene (Nygard *et al.*, 2007).

The RT-PCR products were visualised and assessed using gel electrophoresis (5.2.14). If the PCR products were not clearly visible, the PCRs were run for a further 10 cycles and the products re-assessed by agarose gel electrophoresis. Once the primers were tested by PCR, a qPCR was set up for all the genes for each of the six samples and the control. In order to generate a standard curve, standards were prepared as described in 5.2.15, pooling the six cDNAs (the same as prepared for the PCR) together, and preparing serial dilutions of the pool (1:4, 1:8, 1:16, 1:32, 1:64, and 1:128). The cDNA samples were diluted 1:20 in RNase-free water. The qPCR was run as described in 5.2.15, and the results were analysed with the corresponding software. As indicated previously, the results were checked for possible differences between duplicates or any samples with errors, and then imported into an Excel worksheet. In order to analyse these results with the GeNorm application (Vandesompele *et al.*, 2002), which tests the gene expression stability, the data were transformed following the instruction of the manual and with the help of the example files included in the download of the application.

From the imported results file, the Cts for each gene for the tissue samples, excluding standards and controls, were taken and the duplicates were averaged, to yield the

mean Ct, and then, the standard deviations (SDCt) were calculated. The mean Cts were transformed to relative quantities in relation to the sample with the highest expression level or the lowest Ct value, which was set to 1. For this transformation the amplification efficiency resulting from the standard curve was used in the following formula to calculate the sample quantity (Q).

$$Q = E^{(\mathrm{minCt}-\mathrm{sampleCt})} = E^{\mathrm{deltaCt}}$$

In this equation, E was the transformed amplification efficiency, where 2 was equivalent to a 100% amplification efficiency calculated by the software, the minCt was equal to the lowest Ct value or the mean Ct for the sample with highest expression, which was used to calculate the value of Q for all the samples, and the sample Ct was the mean Ct for the sample for which Q had been calculated. The standard deviation of the Quantity (SDQ) was also calculated as indicated in the following formula.

$$\mathrm{SDQ} = E^{\mathrm{deltaCt}} \mathrm{x}\ \mathrm{LnE}\ \mathrm{x}\ \mathrm{SDsampleCt}$$

where LnE was the natural logarithm of the transformed amplification efficiency, and SDsampleCt was the standard deviation of the Ct values of the samples duplicates, calculated previously. The Q values were calculated for all the samples for each gene, then these values were used to construct a table in Excel, where the first column contained the name of the samples and following columns the corresponding Q values for each of the genes.

The GeNorm application was opened, and the results were loaded. GeNorm defined the internal control gene-stability measure M as the average pair-wise variation of a particular gene with all other control genes. Genes with the lowest M values had the most stable expression. After this first calculation, a stepwise exclusion of the gene with the highest M value (or the least stable) was performed, resulting in a number of genes that qualify as internal control genes due to the stability of their expression level in the samples of interest. To determine the number of genes needed in this study, with the same application, the pair-wise variation $V_{n/n+1}$ was calculated

between the two sequential normalisation factors (NF_n and NF_{n+1}, where n is the number of genes) for all samples within the same tissue panel. A large variation meant that the added gene had a significant contribution to the newly calculated normalisation factor and should preferably be included in order to calculate a reliable normalisation factor. The cut-off for the *V* value was proposed at 0.15, below which the inclusion of an additional internal control gene was not required (Vandesompele *et al.*, 2002).

The calculation of this factor was not necessary at this stage of the study, but it was done in order to understand the mechanisms. Once the genes were selected, which was the objective of this experiment, a normalisation factor would be calculated. As shown in the following formula, the geometric mean of the three genes chosen was calculated for each sample using the quantity values. The use of the geometric mean allowed controlling for possible outlier values and abundance differences among the different genes.

$$NFn = \sqrt[n]{\mathrm{REF1} + \mathrm{REF2} + \mathrm{REF3}}$$

These normalisation factors were rescaled by GeNorm, dividing each one by the geometric mean of all the normalisation factors, in order to distribute the normalisation factors around value 1. With this value the normalised expression level for the gene of interest, *SPP1,* was calculated.

$$\text{Normalised } SPP1 = \frac{\text{Raw } SPP1 \text{ Q}}{\text{Normalisation Factor}}$$

As mentioned previously, normalised *SPP1* calculation was not necessary in this experiment, but in another case the normalised *SPP1* values would be analysed for expression level assessments.

5.3. Results

5.3.1. Animal information and tissue collection

The tissues collected from the first two pigs were preserved in Methacarn, O.C.T or Bouin's or by snap freezing in liquid nitrogen. The quality of the mRNA and the histological structure of these tissues were analysed to assess the different methods of preservation. Taking into account the results of these analyses, tissues from the last seven pigs were preserved in RNA*later*, RNA-Bee, and snap frozen in liquid nitrogen, and fixed in Methacarn. The precautions during the collection were high, keeping all the instruments clean with RNase*Zap*, as well as the surface used during the collection, as described in 5.2.1.

	Pig 1	Pig 2		Pig 3	Pig 4	Pig 5	Pig 6	Pig 7	Pig 8	Pig 9
Liquid nitrogen	✓	✓		✓	✓	✓	✓	✓	✓	✓
O.C.T.	✓	✓		-	-	-	-	-	-	-
Methacarn	✓	✓	⇨	✓	✓	✓	✓	✓	✓	✓
Bouin's	✓	✓		-	-	-	-	-	-	-
RNA*later*	-	-		✓	✓	✓	✓	✓	✓	✓
RNA-Bee	-	-		✓	✓	✓	✓	✓	✓	✓

Table 5.6 Tables summarising the preservation and fixation methods used in the nine pigs.

5.3.2. Validation of preservation of mRNA, integrity and tissue structure with tissue collected from pigs 1 and 2

5.3.2.1. RNA isolation

RNA was isolated from the first pig from all tissue and preservation method combinations.

	Pig 1 - 574			Pig 2 - 509		
	ng/µl	**OD 260/280**	**OD 260/230**	**ng/µl**	**OD 260/280**	**OD 260/230**
Liquid Nitrogen	110.63 - 1133.51	1.75 - 2.11	0.62 - 1.78	77.73 - 1947.51	1.91 - 2.33	0.60 - 1.88
Bouin's	7.24 - 138.38	1.88 - 2.62	0.30 - 1.81	35.57 - 74.52	0.88 - 1.94	-1.42 - 0.19
Methacarn	26.69 - 424.48	2.09 - 2.34	0.87 - 2.03	19.21 - 90.7	1.94 - 2.19	0.32 - 1.30
O.C.T.	15.79 - 39.31	1.52 - 1.77	0.26 - 0.57	37.06 - 154.64	1.95 - 2.11	0.37 - 1.08

Table 5.7 Summary of Nanodrop results from endometrium, placenta and whole utero-placental unit samples for pig 1 and 2 for the four preservation methods used in these collections. The Table shows the range concentrations of the RNA in ng/µl, and the OD 260/280 and OD 260/230 ratios from the spectrophotometry.

The snap frozen in liquid nitrogen and O.C.T. preserved samples were processed with the RNA extraction protocol with RNA-Bee (5.2.6 - Method 1), whereas the Bouin's and Methacarn fixed samples were processed with the RecoverAll™ Total Nucleic Acid Isolation, optimised for FFPE Samples kit (5.2.5). All the RNA samples were quantified (5.2.3). The results for the RNA isolated from liquid nitrogen preserved placenta were not satisfactory, therefore the samples were cleaned with an RNeasy column (5.2.7) and quantification was repeated. The quality of the RNA isolated from samples snap frozen in liquid nitrogen (for placenta only samples cleaned with the column were run), Bouin's and Methacarn samples was evaluated by gel electrophoresis. These processes were repeated with the tissues collected from the second pig, with the only difference being that the liquid nitrogen samples were checked by spectrophotometry and on Agilent Bioanalysers. The results are summarised in Tables 5.7 and 5.8, respectively. The RINs were compared. There were no overall differences between the two pigs for these samples. The resulting RNA concentration from all the isolations (as well as OD ratios) varied between the different preservation methods.

Tissue	Pig 1 Liq nit	Pig 2 Liq nit	Pig 1 Bouin's	Pig 1 Methacarn
Endometrium normal	2.1	2.5	5	2.3
Endometrium smallest	2.2	2.2	2	2.2
Placenta normal	-	-	2	2.8
Placenta small	-	-	2.1	2.5
Whole uterus normal	2.1	2.5	1.8	2.9
Whole uterus small	2.6	2.7	2	2.5
Placenta normal Column	2.1	2.2	-	-
Placenta small Column	2.1	2.3	-	-

Table 5.8 RNA integrity as assessed by the Agilent Bioanalyser. The RNA integrity number (RIN) is shown for the RNA samples from pig 1 for liquid nitrogen, Bouin's and Methacarn and from pig 2 for liquid nitrogen (Liq nit).

The low RIN values (see Table 5.8) indicate high levels of RNA degradation in these samples. These values, together with the absence of the expected 18S and 28S bands, and the spectrophotometry results, demonstrated that the RNA suffered damage.

RNA isolation	Pig 1 - 574			Pig 2 - 509		
	Concentration	OD ratios	RIN	Concentration	OD ratios	RIN
Liquid nitrogen	✓	✓×	×	✓	✓×	×
Bouin's	✓×	✓×	×	✓	×	-
Methacarn	✓	✓×	×	✓×	✓×	-
O.C.T.	×	×	-	✓	✓×	-

Table 5.9 Summary of results for the RNA isolation from pig 1 and 2 for the four different preservation methods with evaluation as defined previously (Error! Reference source not found. **and** Error! Reference source not found.**).** Quality indicators: ✓, acceptable; ✓×, low; ×, none.

5.3.2.2. PCRs and gel results

cDNA was prepared from 48 samples, from pigs 1 and 2 with the RNA analysed in the previous step and PCRs were run for *SPP1*, the gene of interest, and *ACTB*, the chosen internal control gene (5.2.12 and 5.2.13). Figures 5.3 to 5.6 show gel electrophoresis images of the PCR products generated from the RNA isolated from pig 1, for the six tissues collected in the four different preservation methods for *SPP1* and *ACTB*. It has to be noted that for these PCRs the taq/taqstart mixture that help minimise the formation of primer-dimers was not used and evidence of primer-dimers can be seen in the gels for the PCR products for *SPP1*.

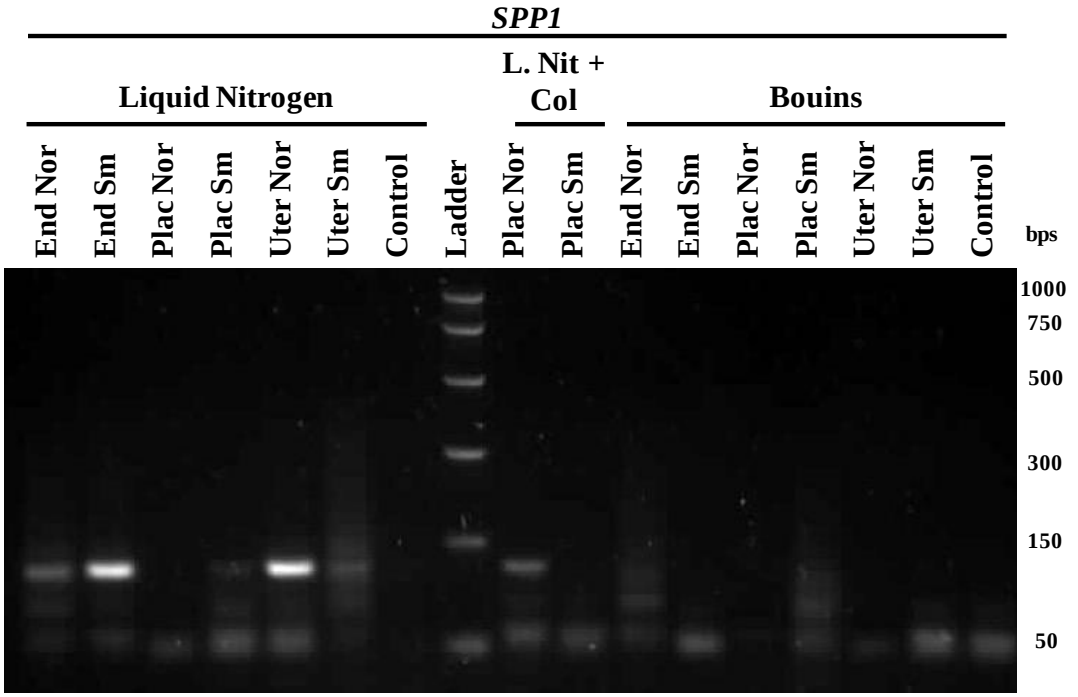

Figure 5.3 Gel electrophoresis results from PCR products for *SPP1* from the RNA isolated from endometrium (End), placenta (Plac) and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 1, preserved in liquid nitrogen and in Bouin's. For liquid nitrogen, the placenta samples were cleaned up with a column and these samples were also used in these experiments (L. Nit + Col). Both sets of samples include the corresponding control for the cDNA construction. A ladder was included to check the size of the amplified fragment for *SPP1*, 120bps.

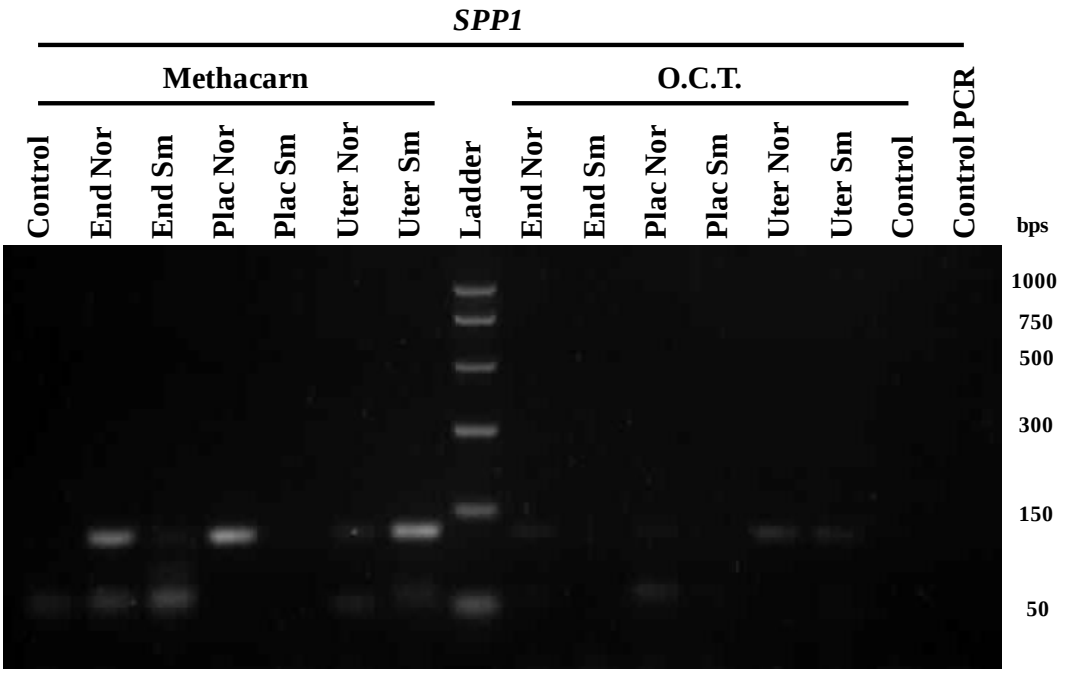

Figure 5.4 Gel electrophoresis results from PCRs products for *SPP1* from the RNA isolated from endometrium (End), placenta (Plac), and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 1, preserved in Methacarn and in O.C.T. Both sets of samples include the corresponding control for the cDNA construction. The lane labelled 'control PCR' shows the absence of products in the *SPP1* PCR control (i.e. without primers). A ladder was included to check the size of the amplified fragment for *SPP1*, 120 bps.

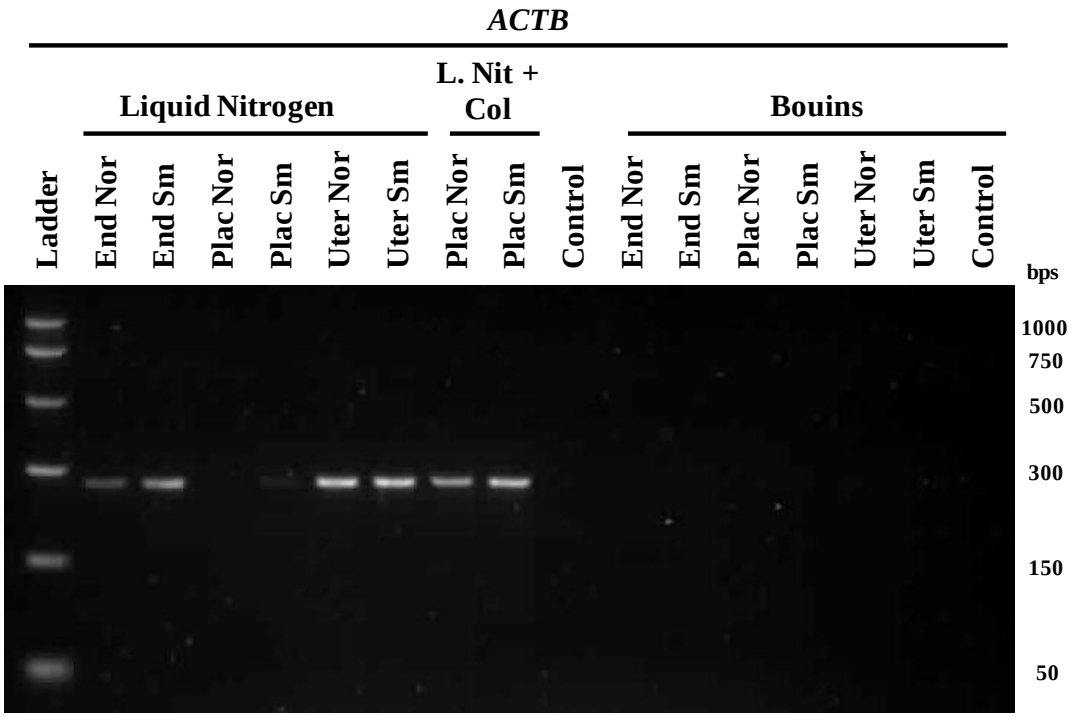

Figure 5.5 Gel electrophoresis results from PCR products for *ACTB* from the RNA isolated from endometrium (End), placenta (Plac), and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 1, preserved in liquid nitrogen and in Bouin's. For liquid nitrogen, the placenta samples were cleaned up with a column and these samples were also used in these experiments (L. Nit + Col). Both sets of samples include the corresponding control for the cDNA construction. A ladder was included to check the size of the amplified fragment for *ACTB*, 259 bps.

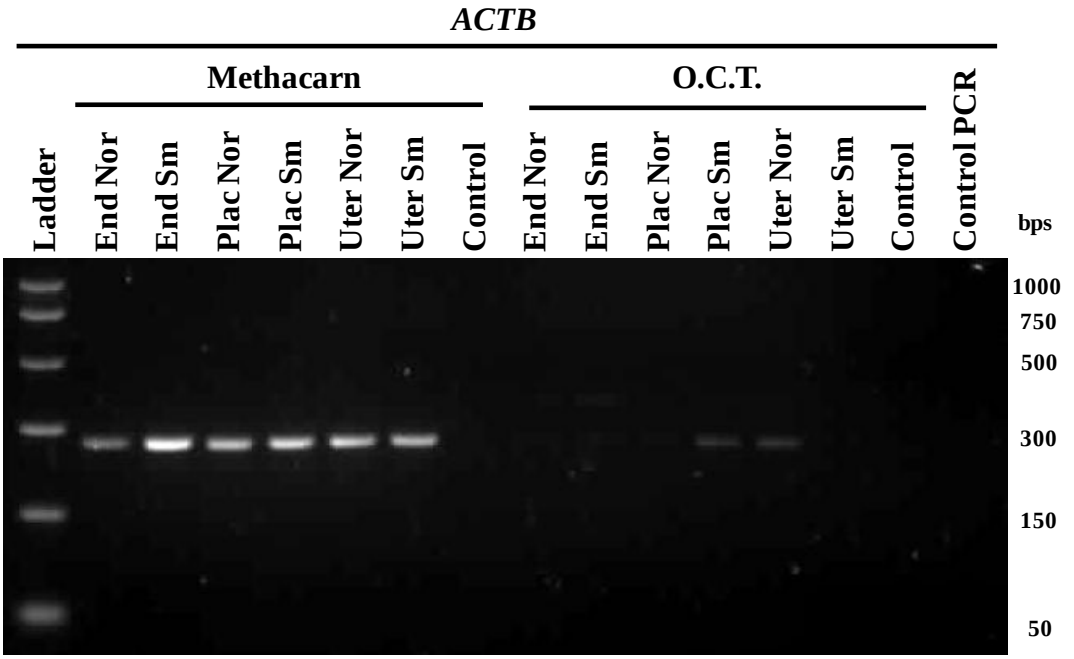

Figure 5.6 Gel electrophoresis results from PCRs products for *ACTB* from the RNA isolated from endometrium (End), placenta (Plac), and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 1, preserved in Methacarn and in O.C.T. Both sets of samples include the corresponding control for the cDNA construction. The lane labelled 'control PCR' shows the absence of products in the *ACTB* PCR control (i.e. without primers). A ladder was included to check the size of the amplified fragment for *ACTB*, 259 bps.

The cDNA (no RT) and PCR (no primers) controls were negative for both genes, confirming the absence of contaminants, including genomic DNA. In respect of *SPP1*, PCR products were poor and weak, or non-existent for the RNA samples isolated from tissues preserved with liquid nitrogen, Methacarn, and O.C.T. There was no evidence of PCR products derived from *SPP1* transcripts in the Bouin's samples. On the other hand, for some treatments it was possible to amplify fragments of *ACTB* transcripts from the corresponding RNA samples. It was possible to detect *ACTB* transcripts in RNA isolated from all the samples preserved with liquid nitrogen or Methacarn, except the placenta samples snap-frozen in liquid nitrogen that had not been cleaned up with the column. No *ACTB* fragments could be amplified from the RNA isolated from the Bouin's fixed tissues. Two weak bands for

placenta small and uterus normal were observed for the O.C.T. samples. Figures 5.7 to 5.10 show gel electrophoresis pictures corresponding to the PCR products for the RNA isolated from pig 2, for the six tissues collected in the four different preservation methods.

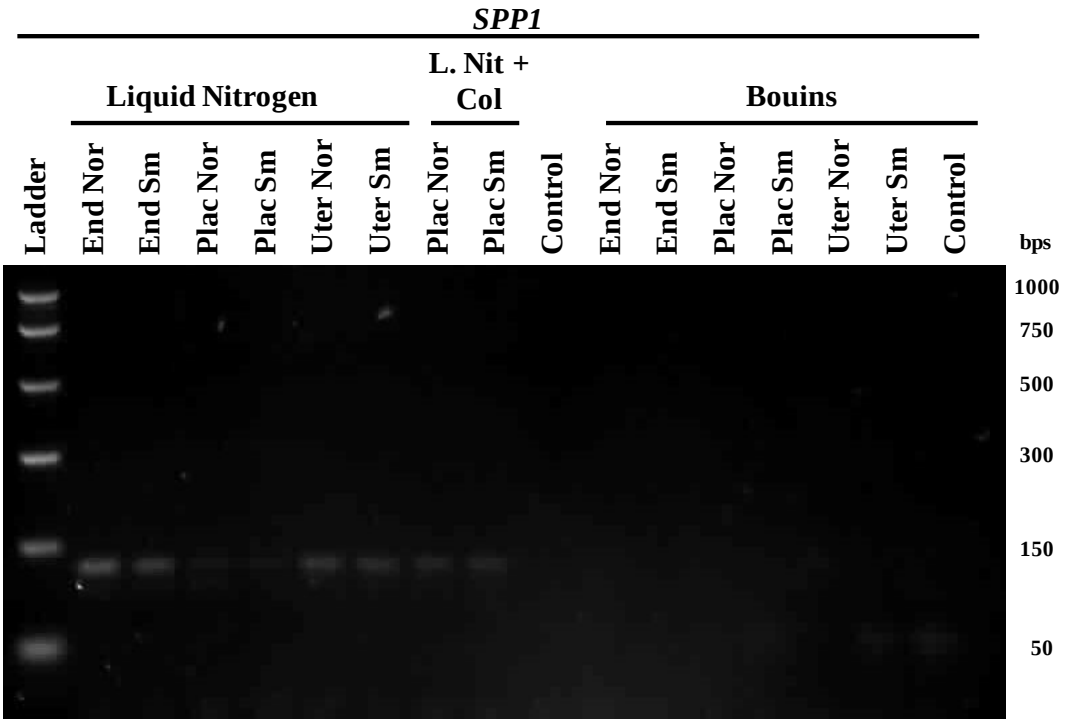

Figure 5.7 Gel electrophoresis results from PCR products for *SPP1* from the RNA isolated from endometrium (End), placenta (Plac) and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 2, preserved in liquid nitrogen and in Bouin's. For liquid nitrogen, the placenta samples were cleaned up with a column and these samples were also used in these experiments (L. Nit + Col). Both sets of samples include the corresponding control for the cDNA construction. A ladder was included to check the size of the amplified fragment for *SPP1*, 120bps.

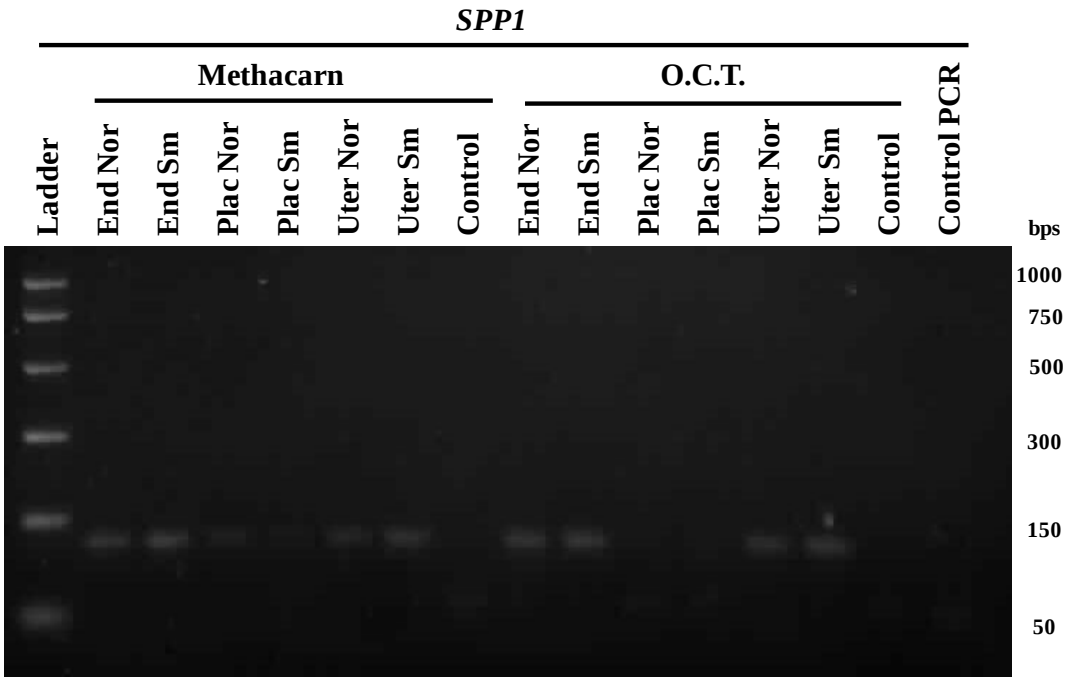

Figure 5.8 Gel electrophoresis results from PCRs products for *SPP1* from the RNA isolated from endometrium (End), placenta (Plac) and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 2, preserved in Methacarn and in O.C.T. Both sets of samples include the corresponding control for the cDNA construction. The lane labelled 'control PCR' shows the absence of products in the *SPP1* PCR control (i.e. without primers). A ladder was included to check the size of the amplified fragment for *SPP1*, 120bps**.**

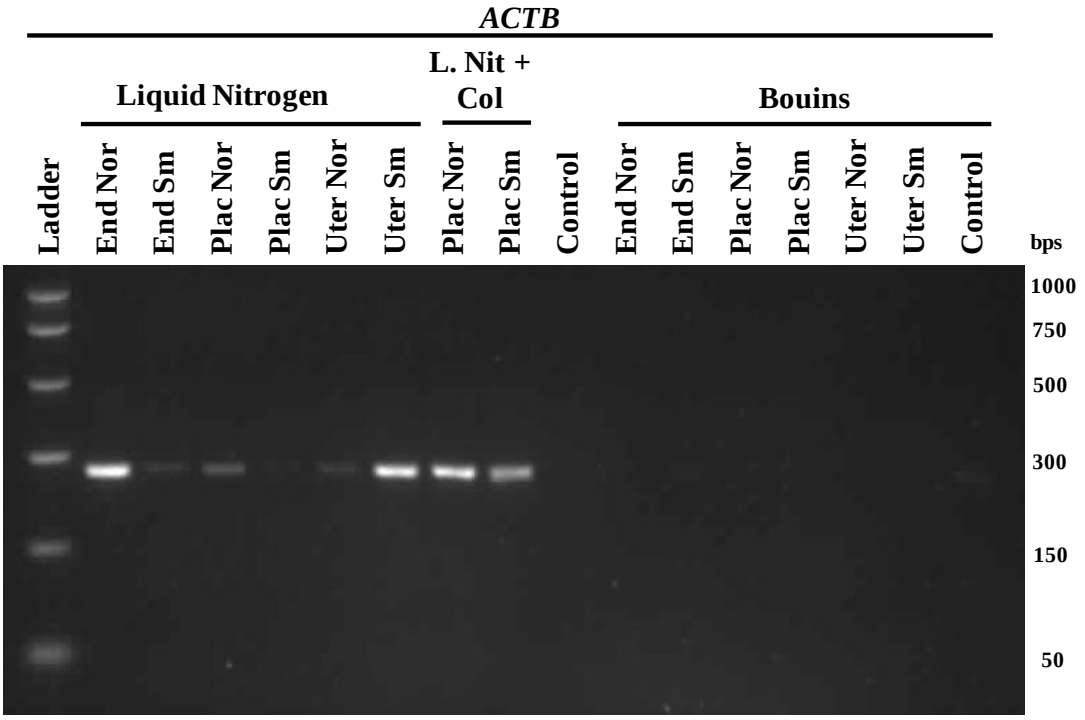

Figure 5.9 Gel electrophoresis results from PCR products for *ACTB* from the RNA isolated from endometrium (End), placenta (Plac) and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 2, preserved in liquid nitrogen and in Bouin's. For liquid nitrogen, the placenta samples were cleaned up with a column and these samples were also used in these experiments (L. Nit + Col). Both sets of samples include the corresponding control for the cDNA construction. A ladder was included to check the size of the amplified fragment for *ACTB*, 259bps.

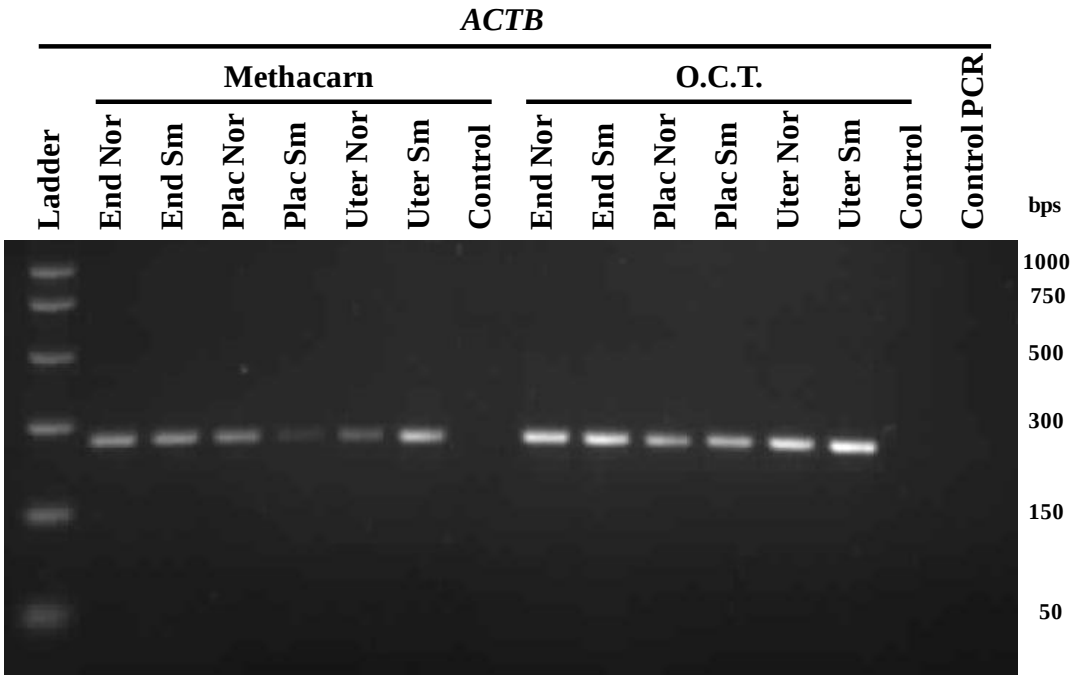

Figure 5.10 Gel electrophoresis results from PCR product for *ACTB* from the RNA isolated from endometrium (End), placenta (Plac) and whole utero-placental unit (Uter), from the smallest (Sm) and from a normal sized (Nor) foetus from pig 2, preserved in Methacarn and in O.C.T. The lane labelled 'control PCR' shows the absence of products in the *ACTB* PCR control (i.e. without primers). A ladder was included to check the size of the amplified fragment for *ACTB*, 259bps.

As for the analysis of RNA from pig 1, the PCR controls (no primers) and cDNA controls (no RT) were negative for both genes. *SPP1* PCR amplification results for pig 2 showed bands for some of the samples, as previously for pig 1. For the liquid nitrogen and Methacarn samples, bands were found for all, with a weaker signal from the placenta sample that was not cleaned up. However, amplified fragments were detected for O.C.T. samples for endometrium and uterus, but not for placenta. Consistent with the results from pig 1, RNA isolated from Bouin's fixed tissue did not show any amplification products. On the other hand, *ACTB* exibited quality results for all the samples, except for Bouin's, where no bands were observed.

5.3.2.3. Haematoxylin and eosin staining comparative study

Haematoxylin and Eosin (H&E) staining for the three tissues fixed in Bouin's, Methacarn, and O.C.T. were examined under a microscope and the preservation of the tissue morphology, presented in Figures 5.11 to 5.13, was compared.

The staining illustrated complete functionality for all the samples analysed, with only differences in colour strength between fixatives. In the endometrium tissue, the glandular and luminal epithelium structures were easily differentiable in the three fixatives (Figure 5.11) without important differences in the tissue histology quality. On the other hand, placenta staining for O.C.T. showed a damaged tissue, whereas tissues preserved in the other two fixatives were intact with a clearer section for the Methacarn fixative (Figure 5.12). Whole utero-placental unit sections (Figure 5.13) illustrated a more intact tissue and clearer pictures for the Methacarn fixative.

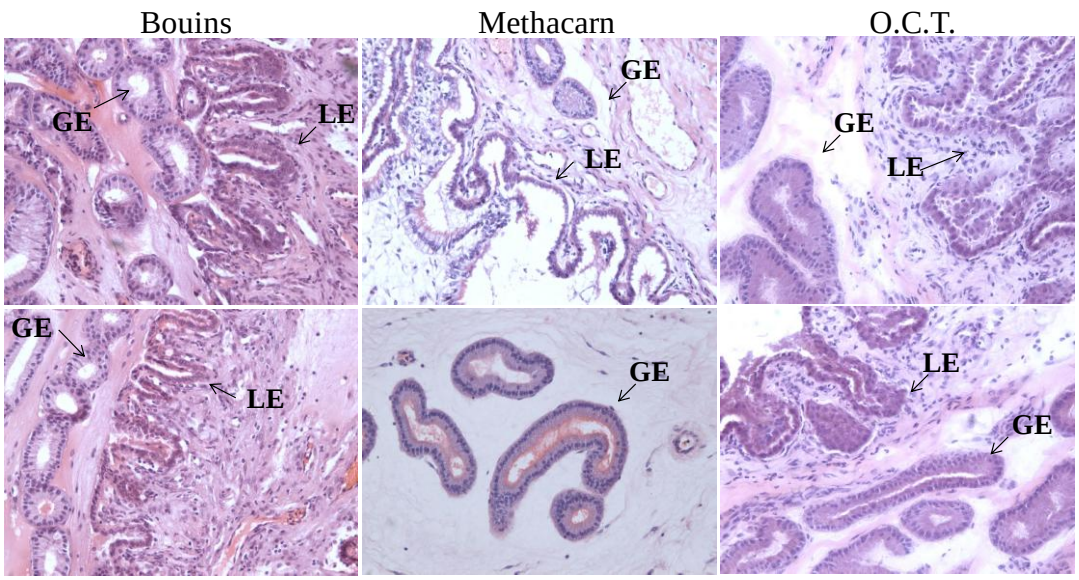

Figure 5.11 Pictures of H&E staining for Endometrium slides from Bouin's, Methacarn and O.C.T. preserved tissues representing the two different structures present in this tissue. Magnification 10x. GE, glandular epithelium, LE, luminal epithelium.

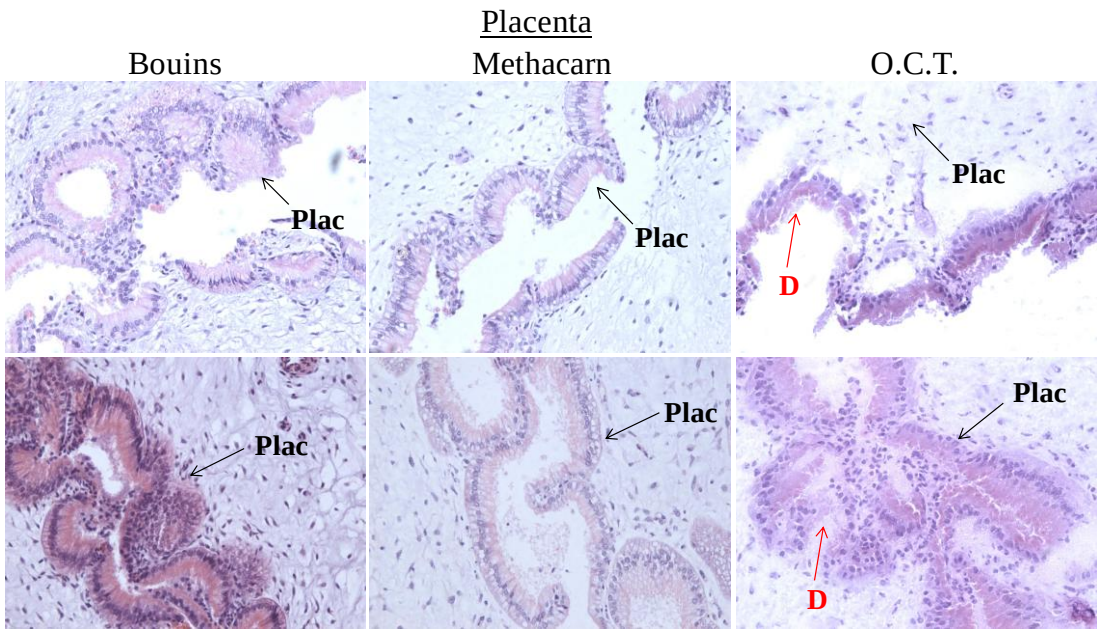

Figure 5.12 Pictures of H&E staining for Placenta slides from Bouin's, Methacarn, and O.C.T. preserved tissues. Plac, placenta, D, damage.

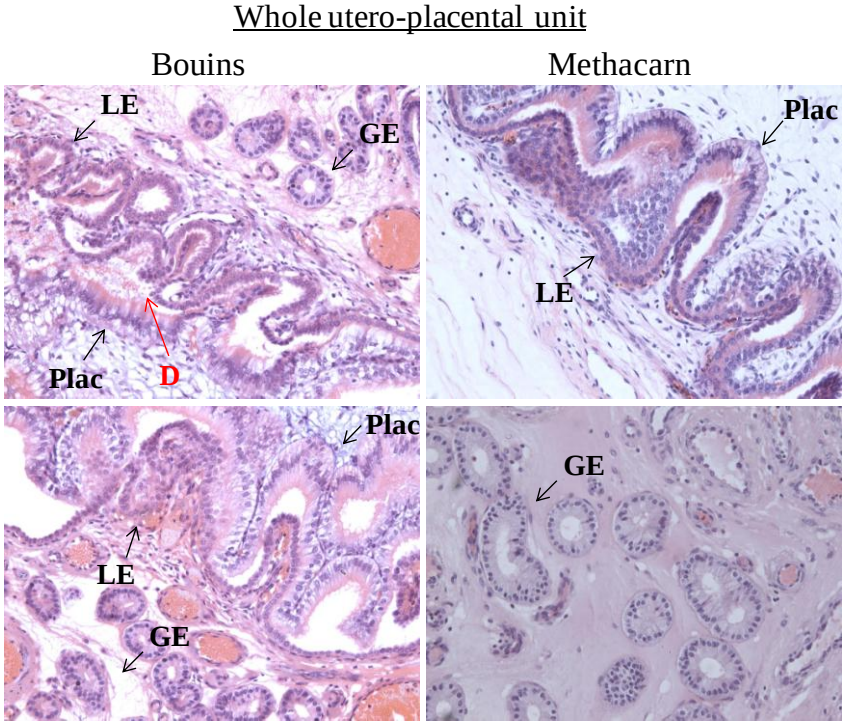

Figure 5.13 Pictures of H&H staining for Whole utero-placental unit slides from Bouin's and Methacarn fixed tissues. GE, glandular epithelium; LE, luminal epithelium; D, damage.

5.3.2.4. qPCR

Quantitative PCR analyses (qPCR) were performed as described in 5.2.12 and 5.2.15 for the RNA isolated from the tissues preserved in liquid nitrogen for the first two pigs. The qPCR was carried out for *SPP1*, with *ACTB* as internal control gene. These analyses were done in order to further assess the quality of the isolated RNA, as well as performance of the qPCR assays for these genes. In this study, qPCR was used to analyse the expression of *SPP1* in the different tissues and in foetus of different sizes.

5.3.2.5. Summary - **Error! Reference source not found.**

The initial assessment of the ability of the different methods for preserving tissue samples to protect the integrity of RNA gave unsatisfactory results. The analyses of RNA integrity using the Agilent Bioanalyser were particularly disappointing and suggested the need for improvements in the sample preservation and/or RNA isolation methods. However, when the isolated RNA was used as a template for RT-PCR for *SPP1* and *ACTB*, the expected fragments were detected, confirming the presence of mRNA from these genes. Thus, whilst the RNA might not be of sufficient quality for some analyses, it may be sufficient for others, including ISH. The RNA preserved in samples fixed in Methacarn was of sufficient quality for ISH analysis. The effectiveness of the conservation of tissue structure varied between the different fixatives with the best quality results found for tissues fixed in Methacarn.

	Pig 1 - 574				Pig 2 - 509			
Experiments	**Liquid nitrogen**	**Bouin's**	**Methacarn**	**O.C.T.**	**Liquid nitrogen**	**Bouin's**	**Methacarn**	**O.C.T.**
RNA Isolation Method	Method 1	FFPE method	FFPE method	Method 1	Method 1	FFPE method	FFPE method	Method 1
Nanodrop	✓	✗	✓	✗	✓	✗	✓✗	✓✗
Agilent	✗	✗	✗	-	✗	-	-	-
RT	✓	✓	✓	✓	✓	✓	✓	✓
PCR	✓	✓	✓	✓	✓	✓	✓	✓
Gel *SPP1*	✓✗	✗	✓✗	✗	✓✓	✗	✓	✓✗
Gel *ACTB*	✓	✗	✓✓	✗	✓✓	✗	✓✓	✓✓
Sections	-	✓	✓✓	✗	-	✓	✓✓	✓
qPCR	✓	-	-	-	✓	-	-	-

Table 5.10 Summary table of the experiments performed with tissues from pig 1 and 2 for the validation of the optimal methodology for the preservation and fixation of the tissues and quality of results. The Table shows which experiments have been performed in each of the available preservation and fixation methods and the quality of the RNA concentrations, the OD ratios and the RIN values. ✓✓, high quality, ✓, acceptable, ✓✗, low, and ✗ no quality.

5.3.3. RNA problem solving

The quality of the RNA isolated from pigs 1 and 2, as described above was variable and in a number of respects, especially its integrity as assessed with the Agilent Bioanalyser, inadequate for the purpose of functional analyses. A number of possible explanations were considered: 1) operator errors, 2) problems with reagents and plastic-ware, 3) use of RNA isolation methods which were suboptimal for the target tissues, and 4) use of inappropriate methods for collecting and preserving the tissue samples.

Operator errors were eliminated as an explanation when similar results were obtained by another more experienced member of the laboratory, including confirming that the integrity of the RNA was inadequate for the proposed experiments. Thus, the reagents used and/or the methods used to isolate the RNA were considered as potential sources of the problems. Firstly, new stocks of all reagent and consumables were prepared. The most significant change was a change in the tubes used; RNase-free tubes were purchased. Secondly, the method was targeted. Two different techniques were used to extract RNA. The protocol (Method 1) used previously was used with two different tubes: autoclaved tubes as used previously and the new RNase-free tubes. All samples were also subjected to an RNeasy column clean up. Method 2, designed for tissues containing high amounts of lipophilic material, like placenta, was used together with the RNeasy column as described in 5.2.8. RNA quantity and quality was assessed (Table 5.11).

Tissue	Method 1 with RNase free tube	Method 1 with autoclaved tube	Method 2 with RNase free tube
Endometrium smallest	8.9 - ✓✓	8.7 - ✓✓	8.6 - ✓✓
Placenta smallest	2.4 - ×	2.4 - ×	1.1 - ×
Endometrium normal	9.1 - ✓✓	9.2 - ✓✓	8.8 - ✓✓
Placenta normal	8.3 - ✓✓	7.8 - ✓✓	2.3 - ×

Table 5.11 RINs for pig 2 endometrium and placenta with two different methods and with two different tubes for one of the methods. The Table indicates the quality of the samples. ✓✓, high quality; ✓, acceptable; ✓×, low; ×, no quality.

The RNA quality was similar for both methods and tubes. Bearing in mind the results shown in Table 5.11, the rest of the tissues from this pig (2) were processed and the placenta small was repeated using Method 1 with RNase-free tubes. These isolations were assessed for the quality before proceeding with the other samples. The results were generally good for Method 1 with RNase-free tubes. Thus, the RNA isolations were repeated using this same method for the tissue samples collected from pig 1. The quality of these samples was assessed, and the results are presented in Table 5.12. The results were not optimal for all the samples.

Sample ID pig 1	Yield (ng/µl)	OD ratios	RIN
Endometrium normal	✓	✓✓	6.3 - ✓
Endometrium smallest	✓	✓×	6.3 - ✓
Uterus normal	✓	✓✓	6.8 - ✓
Uterus smallest	✓	✓✓	3.2 - ×
Placenta normal	✓	×	2.4 - ×
Placenta smallest	✓	×	2.4 - ×

Table 5.12 Summary of results for RNA isolated from liquid nitrogen samples from pig 1 using Method 1 with RNase-free tubes. The Table shows ids of samples and the acceptability or not of the RNA concentrations, the OD ratios, and the RIN values. The Table indicates the quality of the samples; ✓✓, high quality; ✓, acceptable; ✓×, low; ×, no quality.

Meanwhile, further samples were collected from the other seven pigs. RNA*later* was incorporated into the preservation methods due to the better results in yields compared with liquid nitrogen and other techniques as reported by others (Florell *et al.*, 2001; Srinivasan *et al.*, 2002). Before proceeding to isolate RNA from all the tissues, the two new techniques used in this collection, RNA*later* and RNA-Bee, were tested isolating RNA from some additional samples, and assessing its quality and the yields. The RNA isolated from RNA-Bee preserved samples was inadequate. RNA isolated from RNA*later* tissues resulted in appropriate yields, concentrations, and OD ratios, but the RINs were low.

Bearing in mind the problem with the RNA quality from placenta samples, two other RNA isolation reagents and a new method were tested. The new method (5.2.9), similar to Method 1 but with some extra steps to increase the quality of the RNA, was performed on extra samples of placenta tissue preserved in RNA*later* and in liquid nitrogen, using different reagents: RNA-Bee (used up to now), TRIzol (5.2.10), and Ultraspec II isolation reagent (5.2.11). This allowed the comparison not only of the three different reagents with the same method, but also the comparison of two different preservation methods. The results for these isolations are presented in Table 5.13. As reflected in the results, the new method with RNA-Bee (or Method 3) gave better results for both preservation methods than previous methods. Method 4, with TRIzol, also gave adequate results and Method 5, with Ultraspec, gave unsatisfactory results for both tissues.

	Sample ID	ng/µl	OD ratios	RIN
RNA later	**Method 3**	✓	✓×	6.7 - ✓
	Method 4	✓	✓	5.6 - ✓
	Method 5	✓	×	2.9 - ×
Liquid nitrogen	**Method 3**	✓	✓×	6.6 - ✓
	Method 4	✓	✓	6.6 - ✓
	Method 5	✓	✓×	2.4 - ×

Table 5.13 Summary of results for RNA isolated from placenta tissue preserved in RNA*later* and in liquid nitrogen. RNA isolated with Method 3 using three different isolation reagents. The Table shows ids of samples and the acceptance or not of the RNA concentrations, the OD ratios, and the RIN values. The table indicates the quality of the samples; ✓✓, high quality; ✓, acceptable; ✓×, low; ×, no quality.

On the basis of these results, Method 3 with RNA-Bee was used for subsequent isolations of RNA from the placental samples, and the other tissues were processed with Method 1. RNA was isolated from some newly collected tissues from pig 3 and 4 from the samples preserved in RNA*later* and its quality was evaluated, finding lower quality in uterus samples compared with the other studies. Therefore, Method 3 was tried with uterus samples and in view of the improvement Method 3 was

chosen for all isolations. In Table 5.14 results for Method 1 and Method 3 are compared for some tissues, illustrating a general improvement in most of the samples.

Sample ID	Yield (ng/µl)	OD ratios	RIN Method 1+ new tubes	RIN Method 3
Endometrium normal Y22	✓	✓✓	7.9 - ✓✓	-
	✓	✓✓	-	8.2 - ✓✓
Uterus normal Y22	✓	✓×	7.6 - ✓✓	-
	✓	✓✓	-	7.6 - ✓✓
Uterus smallest Y22	✓	✓×	8 - ✓✓	-
	✓	✓	-	8.3 - ✓✓
Placenta normal Y22	✓	✓✓	-	8.4 - ✓✓
Uterus normal Y24	✓	✓×	4.4 - ×	-
	✓	✓✓	-	7 - ✓✓
Uterus normal Y26	✓	✓	-	8.7 - ✓✓

Table 5.14 Results from random samples collected from the 7 later pigs for the Nanodrop and Agilent. Comparison of results from Method 1 and Method 3 RNA isolations. The table shows ids of samples and the acceptability or not of the RNA concentrations, the OD ratios, and the RIN values. The table indicates the quality of the samples; ✓✓, high quality; ✓, acceptable; ✓×, low; ×, no quality.

5.3.3.1. Summary – Table 5.15

Although satisfactory quality RNA isolation was achieved for some of the liquid nitrogen preserved tissues, when some of the isolations were repeated with the same conditions, the results for one of the tissues, placenta, were consistently inadequate. These results, even when improved with RNA*later* samples from the later sample collection, were unacceptable for the planned analysis (RT-qPCR) and therefore another method was adopted for placenta tissues. The results after introducing this method were excellent, and the results for other tissues using this method were examined, finding superior quality in most cases. As a result, this isolation method was adopted as the method for all the isolations in this study. The comparative analysis of the results with the different methods is summarised in Table 5.15.

	Pig 1	Pig 2	Extra samples			Pig 3	Pig 4	Pig5
			Liquid nitrogen	RNA*later*	RNA-Bee			
Method 1	×	×	-	-	-	-	-	-
Method 1 + RNase free tubes	✓×	✓✓	-	×	×	×	✓×	-
Method 2	-	✓×	-	-	-	-	-	-
Method 3	-	-	Plac ✓✓	Plac ✓✓	-	✓✓	✓✓	✓✓
Method 4	-	-	Plac ✓	Plac ✓	-	-	--	--
Method 5	-	-	Plac ×	Plac ×	-	-	-	-

Table 5.15 Summary table of the RNA isolations in preparation of the optimal method. The Table shows the methods used, the samples used in each method and the acceptability or not of the RNA concentrations, the OD ratios and the RIN values. The table indicates the quality of the samples; ✓✓, high quality; ✓, acceptable; ✓×, low; ×, no quality. Highlighted samples indicate the method with the best results.

5.3.4. Selection of internal control genes for the qPCR analysis

As mentioned before, for the quantification of mRNA level in different tissue samples, and to be able to compare them, an accurate and adequate normalisation for the tissues to be analysed was essential. In order to decide which genes were the best internal control genes, cDNA samples (5.2.12) prepared from the RNA isolated from pig tissues preserved by snap-freezing in liquid nitrogen, were used for qPCR analyses with nine different potential internal control genes. The optimal and working conditions of the primers for these genes were tested through PCRs first for each primer (5.2.13). The PCR amplification was performed for 30 cycles with an Tm of 60ºC. The products were loaded in a gel as described in 5.2.14, and the results are presented in Figures 5.16 to 5.18. Results from some of the primers in this initial PCR are not presented, due to the absence of results. The results for the *ACTB*, *B2M1*, and *YWHAZ* genes were satisfactory (Figures 5.14 and 5.15).

The PCR assays for the other genes (*GAPDH*, *HMBS2*, *HPRT1*, *RPL4*, *SDHA*, and *TBP1)* were run for 10 more cycles (i.e. a total of 40 cycles) in order to increase the amount of DNA amplified and be able to visualise it, and the products examined by gel electrophoresis, as presented in Figures 5.16 to 5.18. In the last gel, *SPP1* was also loaded to see the product results from the previous PCR, with only 30 cycles.

After these extra cycles, all the genes illustrated clear and well defined bands of the expected sizes in the gels (Figures 5.16 to 5.18), proving that all the primers were working and compatible with the tissues. qPCRs for these genes were prepared as described in 5.2.15. The standards and the dilutions of the samples (1:20) were prepared and the analysis was run as described in 5.2.18.

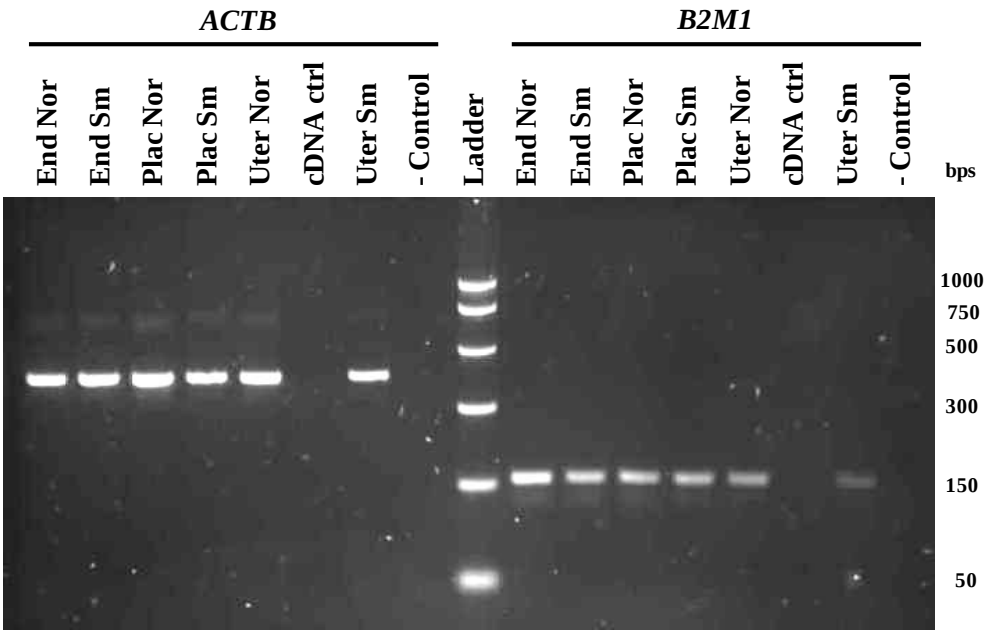

Figure 5.14 Gel electrophoresis results from PCR products for six different RNA samples from pig 2 (liquid nitrogen preserved samples) for primer pairs for *ACTB* and *B2M1*, including cDNA and negative controls, and a size ladder. (bps). End, endometrium; Plac, placenta; Uter, whole utero; Sm, smallest; Nor, normal sized.

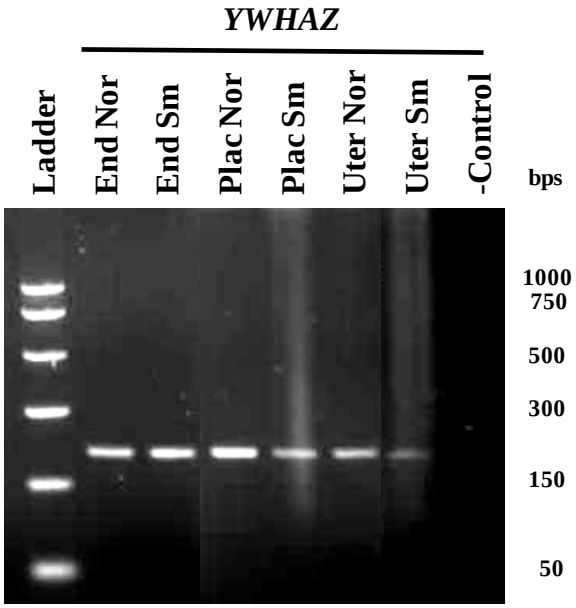

Figure 5.15 Gel electrophoresis results from PCRs products for six different RNA samples from pig 2 samples for primer pairs for *YWHAZ*, including negative control, and a size ladder. End, endometrium; Plac, placenta; Uter, whole utero; Sm, smallest; Nor, normal sized.

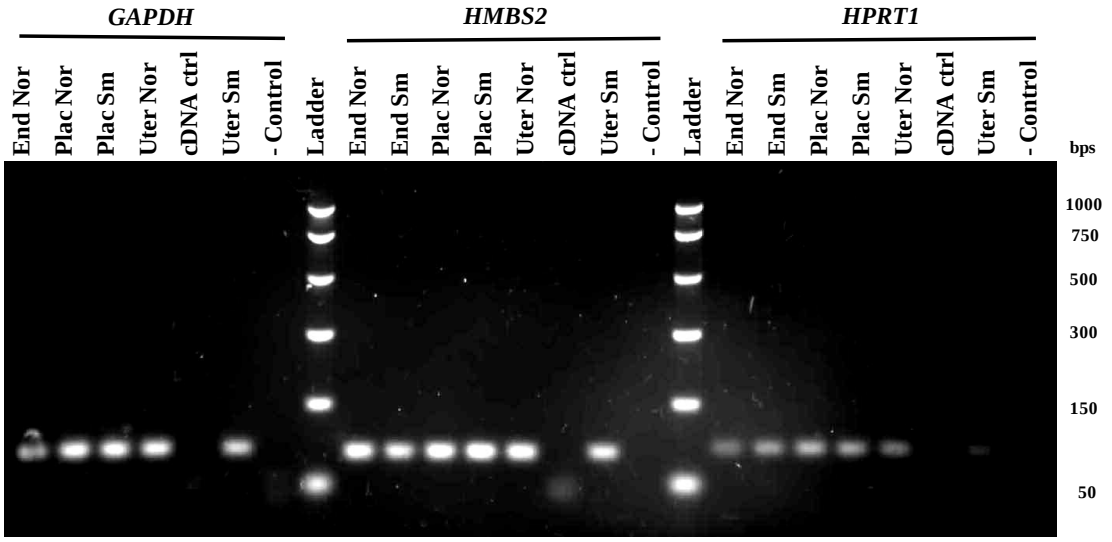

Figure 5.16 Gel electrophoresis results from PCR products for six different RNA samples from pig 2 samples for primer pairs for *GAPDH*, *HMBS2* and *HPRT1*, after 10 extra cycles, including negative control and a ladder. End, endometrium; Plac, placenta; Uter, whole utero; Sm, smallest; Nor, normal sized.

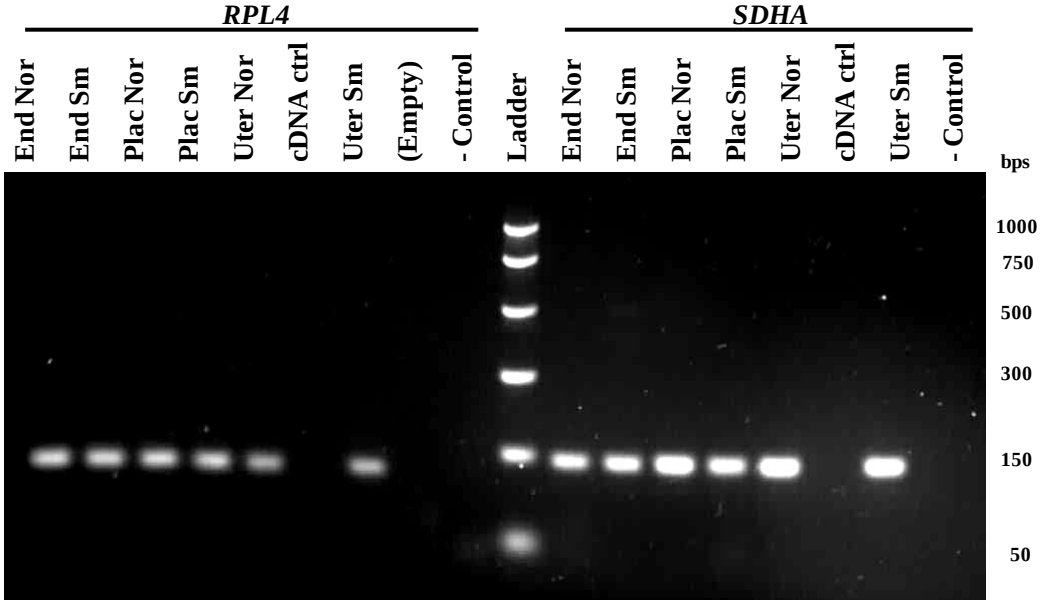

Figure 5.17 Gel electrophoresis results from PCRs products for six different RNA samples from pig 2 samples for primer pairs for *RPL4* and *SDHA*, after 10 extra cycles, including negative control and a size ladder. End, endometrium; Plac, placenta; Uter, whole utero; Sm, smallest; Nor, normal sized.

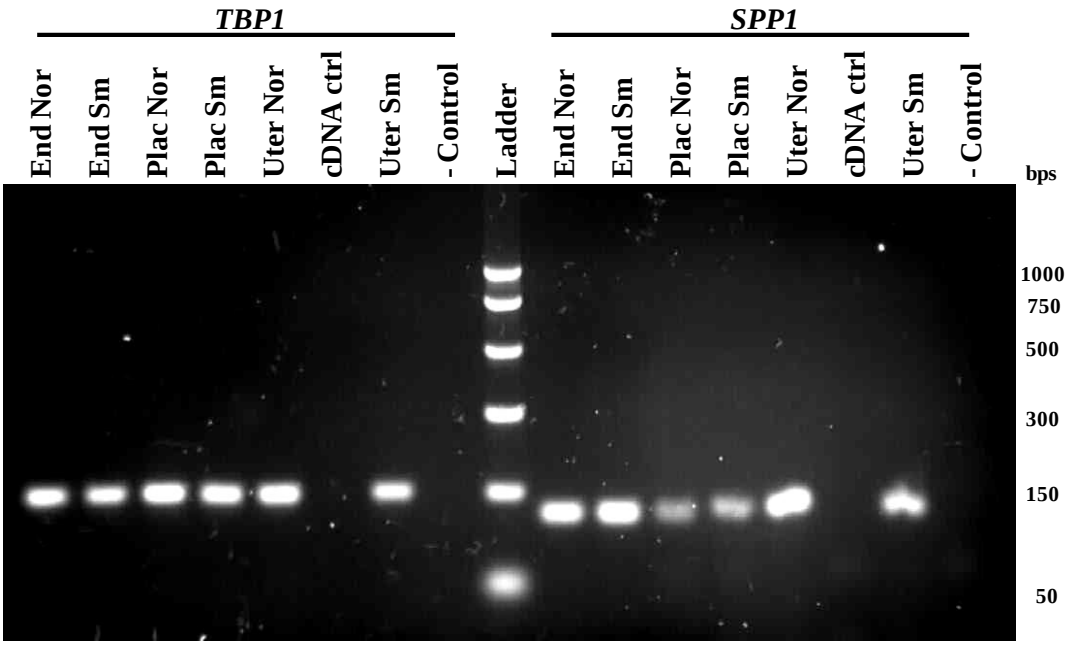

Figure 5.18 Gel electrophoresis results from PCRs products for six different RNA samples from pig 2 samples for primer pairs *TBP1*, after 10 extra cycles, and *SPP1*, including negative control and a size ladder. End, endometrium; Plac, placenta; Uter, whole utero; Sm, smallest; Nor, normal sized.

The results for all the genes were checked for consistency between duplicates, and the linear correlation coefficients (R^2) and the E were also inspected for the quality of the qPCRs for each gene (Table 5.16). Although the results for two of the genes, *ACTB* and *SDHA*, were unsatisfactory, all the results were included in the GeNorm analysis, through which the inadequate internal controls can be identified and excluded. The linear correlation coefficient of the seven genes left and *SPP1* ranged from 0.966 to 0.997 (the optimum is around 1) and the amplification efficiencies for the standard curve calculated by the software ranged from 74.1% to 100.9% (the optimum is between 90 and 110%). The Ct values for the nine internal control genes in all the samples were within 19.36 to 37.12 cycles, and all the samples were within the standard curve resulting from the serial dilutions of the sample pool.

Gene name	R^2	E (%)
ACTB	0.222	176.6
B2M1	0.966	100.3
GAPDH	0.996	87.1
HMBS2	0.994	95.6
HPRT1	0.987	80.6
RPL4	0.983	74.1
SDHA	0.001	>1000
TBP1	0.995	100.9
YWHAZ	0.997	89.5
SPP1	0.980	95.9

Table 5.16 qPCR results for the nine internal control genes examined plus *SPP1*. R^2 = linear correlation coefficient for each qPCR, E= amplification efficiencies of the standards.

The amplification efficiencies were transformed, dividing them by 100 and adding 1 to each one and used to calculate the quantities needed for the GeNorm analyses. The rest of the transformations are indicated in 5.2.18. The expression stability or *M* value of these genes was calculated with the resulting quantities for the nine genes. The least stable, with the highest *M* value (8.095 for *SDHA*), was removed and the calculations were repeated. The next gene with the highest *M* (3.467 for *ACTB*) was removed and the *M* values were recalculated. These steps were repeated until only the most stable genes were left, four in this case, which were *GAPDH*, *HPRT1*, *RPL4, and TBP1* (Table 5.17).

Step	*ACTB*	*B2M1*	*GAPDH*	*HMBS2*	*HPRT1*	*RPL4*	*SDHA*	*TBP1*	*YWHAZ*
1	3.820	2.376	1.856	2.102	1.803	1.827	8.095	1.879	1.893
2	3.467	1.447	0.938	1.220	0.892	0.921	*	0.954	0.974
3	*	1.066	0.514	0.816	0.481	0.513	*	0.557	0.578
4	*	*	0.397	0.682	0.365	0.396	*	0.510	0.521
5	*	*	0.388	*	0.310	0.368	*	0.403	0.415
6	*	*	0.324	*	0.278	0.301	*	0.500	*

Table 5.17 *M* values for all the genes and stepwise exclusion of the least stable gene and recalculation of *M* values (by row).

This stepwise exclusion of genes based on the *M* value is also represented in Figure 5.19. The four genes selected by GeNorm application as the most stable ones had an *M* value lower than 0.5, which was stated as optimal in the manual (GeNorm application manual).

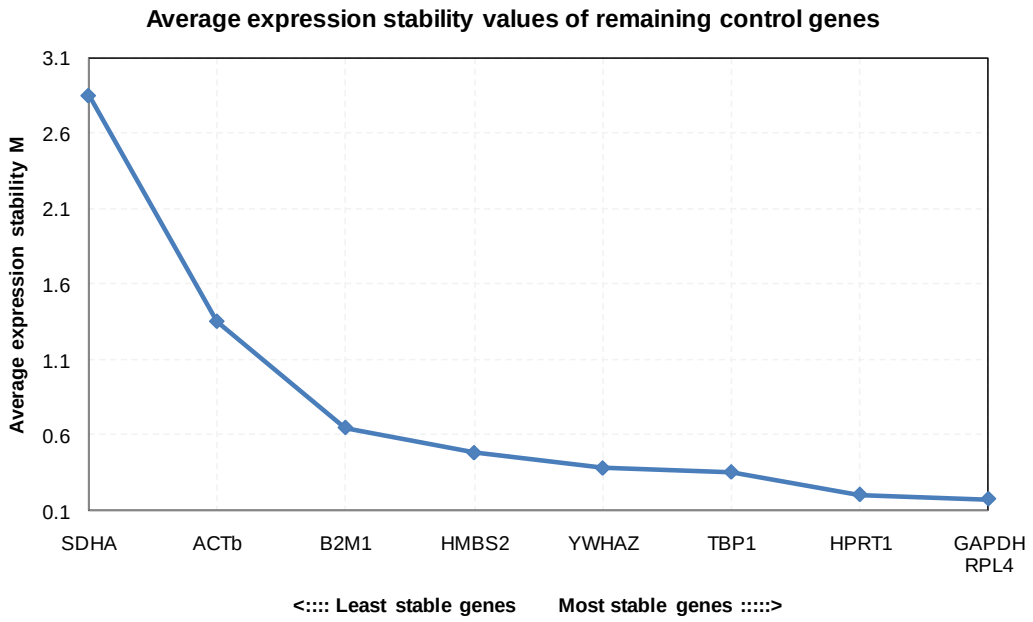

Figure 5.19 M values for the nine potential internal control genes of interest, in order of stability and stepwise exclusion of the least stable.

Once the most stable genes were identified, the *V* value or pair-wise variation was used to determine the optimal number of internal control genes needed for the normalisation. The *V* values, indicated in Figure 5.20, expressed the pair-wise variation between two sequential normalisation factors containing an increasing

number of genes. In this context, variation means that the addition of that gene has a significant effect and it should be included in the calculation of the normalisation factor. In this case, the difference between 3 or 4 control genes was not very large, and the pair-wise variation from 2 to 3 was already lower than 0.15, which had been proposed as cut-off by Vandesompele *et al.* (2002). Moreover, the number of genes to analyse is not large, and in most cases three internal control genes are sufficient. Therefore, *GAPDH*, *RPL4*, and *HPRT1* were chosen as internal control genes for the study of the expression level of *SPP1* in pig reproductive tissues (Vandesompele *et al.*, 2002).

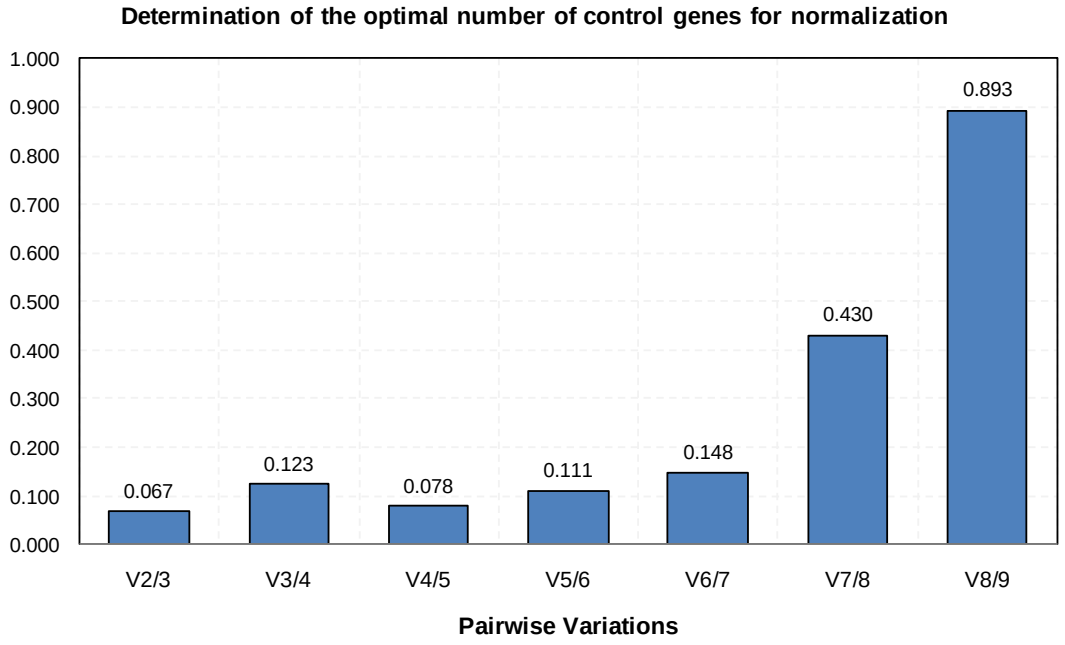

Figure 5.20 Pair-wise variations for the internal control genes showing the effect on the normalisation factor of the addition of another gene in each column.

5.4. Discussion

In order to characterise candidate genes underlying quantitative trait loci with effects on female reproductive traits, different tissues were collected from pregnant gilts/sows using different preservation methods. As mentioned previously, the studies

planned for these tissues involved mRNA and protein analyses of reproductive tissues.

Studying the expression of candidate genes for female reproductive traits presents a number of challenges. Some of the tissues of interest are temporarily formed during pregnancy, such as the placenta, whilst others undergo major transformations during this period. Also the dynamism of all the tissues, proteins and genes implicated in the pregnancy is high with important differences between stages. For these reasons, it is not an easy tissue to collect, and the stage of pregnancy is an important factor to take into account; the gestation period in pigs is ~114 days.

Pigs, compared with mice for example, are a difficult animal from which to collect any tissues. First, the cost of a pig experiment is high, not only due to the cost of the animal, but also the facilities and personal needed to maintain them. Moreover, in order to get tissues from pregnant pigs, as in this case, the pig should reach puberty (> 5 months of age), and have at least another oestrus before mating, which needs to be successful for pregnancy to progress. Once they are pregnant, there is a waiting period until the stage of interest is reached. The collection of the tissues of interest requires surgical procedures, and generally, the sacrifice of the animal. For this, qualified personal need to be present and certain facilities are required.

As noted earlier, pigs are genetically heterogeneous. Thus, it is desirable to perform as many of the different analyses or assays on the same animals, and to sample sufficient animals. The latter impacts on the costs of the experiments, and the former drives a requirement to sample each animal for multiple purposes in a manner that does not compromise the quality of the samples. Given this requirement for multipurpose sampling, the conditions and duration of sample collection are important. Therefore, a certain number of people are required during the collection, in order to achieve the objectives. To this, the time cost and the possible waste due to a bad experiment design need to be added. For these reasons the value of each sample collected in this study is high and its quality is extremely important. These factors drove the extensive optimisation work described in this Chapter, not only the

preservation but also the optimal use of the samples in order to take full advantage of the material collected.

The effects of different preservation and fixation methods on the characteristics of the tissues were compared. First, in order to assess RNA quality, RNA was extracted from all tissues and preservation method combinations. The quality of the isolated RNA was assessed by spectrophotometry, with the Agilent Bioanalyser and the transcription of different internal control genes was examined by PCR, and RT-qPCR. Secondly, to assess tissue structure, H&E staining was performed on all tissues, for all methods. There is no universal method for collecting samples, fixing or preserving them, and proceeding with the particular analysis. All this depends on factors such as the objectives of the study or the material to analyse.

As demonstrated in this Chapter, the isolation of intact RNA from tissues collected from pregnant gilts/sows was challenging. The initial method used for isolation of RNA had been used previously with other tissues with adequate results. However, the results with uterus and placental tissues were not acceptable. The isolation was performed initially from samples preserved with each of the potential preservation methods in order to compare them. However, the snap frozen samples were the only one designated for the RNA expression studies. Initially, the RNA isolated from tissues preserved in liquid nitrogen was of insufficient quality for the proposed analyses. Thus, in this study, measures were required to find an optimal RNA isolation method, due to the need for high quality RNA for subsequent analyses. The use of RNA*later*, which is very effective for protecting RNA, was adopted during these optimisation experiments. After an exhaustive quality test, the RNA isolation methodology was optimised, achieving high quality RNA for the expression level studies through qPCR, with tissues preserved in liquid nitrogen and in RNA*later*. The most challenging tissue for the isolation of quality RNA was the placenta since it has a high lipophilic (McNeil *et al.*, 2007) content and is rich in RNase.

Apart from the need to isolate RNA of high quality and integrity, the major concern was the conservation of tissue structure for histological analysis. Two of the experiments planned, involved the use of histological sections for the location and

quantification of mRNA and proteins in the tissue, ISH and IHC. These experiments require effective conservation of the tissue structure, as well as intact and accessible mRNA and proteins.

In order to check the efficacy of using O.C.T., Bouin's and Methacarn preserved samples for ISH studies, the quality of the RNA present in these tissues was assessed. The result of this assessment demonstrated that the use of Bouin's was inappropriate for ISH on the target tissues. The quality of RNA in O.C.T. treated samples was acceptable. However, RNA from Methacarn fixed samples was better and of the highest quality. The processing including sectioning of the tissue was also simpler for Methacarn treated samples.

Similarly, the results from the histological examination through staining led to the conclusion that Methacarn fixed reproductive tissues were of better histological quality than the other two fixatives analysed in this study. These results together, indicated that Methacarn was the most suitable fixative for the planned analyses.

The tissue integrity was not assessed in snap frozen tissues due to the known aggressiveness of the method, good for preserving RNA and DNA but not for tissue structure due to the rapid freezing and the consequent formation of destructive ice crystals. During the snap freezing process, a very quick way of freezing tissues, there is damage to some extent to the tissue structure. Thus, tissues preserved in this way are not optimal for histological studies. The equivalent method would be the O.C.T. which involves a slower freezing using also liquid nitrogen but the tissue is embedded in O.C.T., a mountant. The tissue would be ready for sectioning, however, with this technique, the cryosectioning, requires more practice and as it was demonstrated by the H&E staining the quality of the tissue structure was not adequate.

The tissues preserved in RNA*later* were not assessed for tissue integrity in this study, but other studies have demonstrated the possibility of retrieving good quality sections (Srinivasan *et al.*, 2002). For this retrieval, the tissue needs to be cleaned up to remove the RNA*later* and then fixed as fresh tissue. If the RNA integrity is also

important the washes need to be done carefully as once the RNA*later* is removed the RNA is exposed for degradation.

As stated in the literature, the normalisation process as part of a quantitative study needs to be customised for each study. In this study, genes, which had been earlier used by others, were first confirmed as possible internal control genes for the tissues of interest. Secondly, the technique proposed by Vandesompele *et al.* (2002) was used to decide which were the appropriate genes for normalisation, and the number needed. As a result, from the nine genes considered, three of the four stably expressed genes were needed for an accurate normalisation of the samples.

In summary, as a result of this validation process, optimal preservation and fixation methods were defined (liquid nitrogen, RNA*later* and Methacarn) together with a high quality RNA isolation method, and internal control genes were selected for a normalisation process for these particular tissues. Also the expression of *SPP1* mRNA in these tissues at this stage of pregnancy was confirmed.

Chapter 6
Functional analysis of SPP1 mRNA and protein in reproductive tissues from pregnant gilts and sows of different breeds

6.1. Introduction

As mentioned previously, the ultimate goal of a QTL study is the discovery of a positional and physiological candidate gene in a region, with a function in the trait of interest. In the analysis performed in the present study and described in Chapter 3, a positional and physiological candidate gene was found, *SPP1*. *SPP1* is expressed in porcine reproductive tissues including placental and endometrial tissue, presenting different patterns of expression throughout the stages of pregnancy (Johnson *et al.*, 2009).

In this study it was hypothesised that differences in foetal growth may be associated with the effectiveness of conceptus attachment, as measured by SPP1 expression. SPP1 has been shown to have a function in implantation and maintenance of pregnancy. The position and function of *SPP1*, together with the hypothesis, set the objective of the work described in this Chapter. Thus, the aim of this study was to characterise SPP1 between feto-placental units of different size occupying the same uterus and in different genotypes with clear differences in LS. In order to fulfil this, placenta, endometrium, and whole utero-placental units were collected from pregnant gilts/sows at day 40 to 45 of pregnancy. This particular stage was chosen because 1) the difference between small and normal foetus was established and detectable and 2) the integrity of uterine and placental tissues allowed clear recognition of these tissues and their collection. These tissues were collected from the feto-placental unit surrounding the smallest foetus of the litter, and from one other foetus of average or normal weight. The characterisation of SPP1 in this study consisted of mRNA location and quantification by qPCR and *in situ* hybridisation (ISH) and protein location and quantification by immunohistochemistry (IHC).

6.1.1. Foetal weight and factors implicated

LS is a composite trait determined by different traits, such as OR, ES, PS, UC and foetal development, which need to be considered for the increase of LS. The greatest limitations to reproductive efficiency across mammalian species are embryonic

mortality and prenatal losses (Roberts & Bazer, 1988). Thus, the increase of LS should be defined as an increase in the number of piglets born alive and viable piglets, together with a reduction of embryo and foetal losses.

Pigs experience two periods of significant loss of conceptuses (embryos/foetus and associated membranes). The first losses occur during the attachment phase of the peri-implantation period and involve 30% of conceptuses (Pope, 1994). These losses are characterised by abnormalities and asynchrony between conceptus signals and uterine receptivity, resulting in defective implantation and/or placentation. The second ones involve 10-15% of the remaining conceptuses, and occur during early to mid-gestation (Pope, 1994).

When uterine space becomes limiting, in late stages of pregnancy, the competition between foetuses produces more losses and the reduction in the size of some of the foetuses with a consequent retardation in development. Some of these foetuses may not survive farrowing and others will have an extreme low birth weight. These piglets are defined as runt pigs and can be identified as lying outside developmental average at a very early stage of pregnancy. The weight at birth is an important factor with a large impact in the development and postnatal survival of the piglet (Foxcroft *et al.*, 2006). Thus, runt pigs have a low change of survival to weaning and even to enter the production chain. These low birth or runt pigs can be identified at early stages of pregnancy, as the variation in within-litter birth weight is established by day 35 of gestation; however, the mechanisms by which this occurs, are not fully understood (Foxcroft *et al.*, 2006).

6.1.1.1. Maternal effects

In a reciprocal embryo transfer analysis, the sizes of the foetuses from another breed carried in a MS gilt uterus were smaller than foetuses growth in other uterus (Ashworth *et al.*, 1990b). This suggested a foetal growth limitation due to the MS maternal environment. The different secretion during pregnancy in MS uterus has been suggested to be also a maternal or genotype factor.

6.1.1.2. Litter size and uterine capacity

Litter size is determined by uterine space and nutrient exchange, factors which demand an increase with the advance of the pregnancy. As a consequence, the number of piglets in the uterus is limited by the uterine space and by the increase in placenta size due to the demand of the foetus. Thus, to increase the feto-maternal exchange rate an increase in the placenta size is required. With this increase, the UC is reduced and the uterus is crowded, reducing the growth of some foetuses.

6.1.1.3. Position in the uterus

One factor implicated in fetal weight differences has been considered to be the position of the foetus in the curved porcine uterine horn. In a study of the foetal weight and the vascular supply in the uterine horn by Perry and Rowell (1969), the heavier foetuses were found at the ovarian end of the horn, when the number of foetuses in the horn was higher than 5, with the lighter foetuses in the middle when the number of foetuses was high. In another study, Wise *et al.* (2001) detected no relation between the uterine position and the foetus weight at day 30 of pregnancy. Moreover, Perry and Rowell (1969) and Wise *et al.* (1997) found that at the end of gestation in pigs, heaviest foetuses were located in the ovarian ends, and there was a decline in foetal weight from the ovarian end toward the middle. However, results from other studies (Ashworth *et al.*, 2001; Finch *et al.*, 2002) suggested no relationship between foetal size and position due to a random spacing of the foetuses in the uterus. Beside, porcine embryos have the capability to move and space themselves within the uterus during the early stages of pregnancy (Pope *et al.*, 1982), and the effect of this spacing has an important influence on subsequent growth and development. Therefore, due to the many factors implicated, there is not a predictable position for the lighter embryos or foetus in a litter.

6.1.1.4. Foetal sex

From mid gestation, foetal sex has been associated with an increased size of male foetuses and their associated placenta (Wise *et al.*, 1997). Furthermore, the sex of the neighbouring foetuses was seen to affect foetal size at late gestation.

6.1.1.5. Placental nutrient transport/implantation

As previously mentioned, pigs have a non-invasive, epitheliochorial placentation, where nutrients need to travel an increasing distance from the maternal blood supply to the foetal blood supply through epithelial cells and cell membranes (Friess *et al.*, 1980). Therefore, the efficiency of the placenta is dependent on the contact surface with the uterine wall. Small placentas have been associated with the smallest foetus in the uterus.

6.1.2. Meishan *vs*. other breeds differences in weight, litter size, implantation and placentation

In Chapter 3, *SPP1* was confirmed to be a candidate gene for PS in a QTL study in a LW x MS cross population. For this reason, the study of this gene in both breeds with marked differences in reproductive performance was one of the objectives of this study. As described previously, there are numerous studies on SPP1 mRNA and protein expression in reproductive tissues in pigs and other species, especially in sheep (Zhang *et al.*, 1992; Johnson *et al.*, 1999a; Johnson *et al.*, 1999b; Apparao *et al.*, 2001; Carson *et al.*, 2002; Apparao *et al.*, 2003; Johnson *et al.*, 2003a; Johnson *et al.*, 2003b; Johnson *et al.*, 2003c; White *et al.*, 2005; White *et al.*, 2006; Allan *et al.*, 2007; Erikson *et al.*, 2009). These studies analyse not only the location but also the regulation of SPP1 mRNA and protein in these tissues. However, there is currently no similar study comparing SPP1 differences in size within a litter or between differing breeds.

MS pigs have been shown to have a higher LS than other breeds through a higher level of PS for a given OR (Haley & Lee, 1993). The MS strategy to achieve this higher performance have been assigned to the smaller size of MS conceptus at all stages, the smaller placenta at term, the uniform development of the conceptus and synchrony, mainly at the early stages of pregnancy, and the reduction in uterine secretion during early pregnancy, such as oestrogen (Ford, 1997; Vallet *et al.*, 1998) among others. It is not clear though which ones are the causes of the larger LS and which are the consequence of these larger litter. For example, the smaller size of the foetus will consequently reduce the competition for space in the uterus that is also better distributed between embryos with more equal space between embryos. But maybe it is the equal distribution that made them have less competition, and thus, grow at a similar rate, also a consequence of the reduction in uterine secretion. Also, the smaller size of the placenta has been detected to be due to an increased efficiency and higher blood vessel density. These factors define the efficient strategy of the MS gilts and sows compared to European and US breeds

In a recent study by Fernandez-Rodriguez *et al.* (2011), published during the production of this thesis, *SPP1* gene was identified and validated as differentially expressed in ovaries in a microarray study comparing expression of high and low-prolificacy sows. In a study in the same population (F_2 MS x Iberian) by Noguera *et al.* (2009), an epistatic QTL on SSC8 was mapped for TBA. The confidence interval for TBA was located in a similar position to the one mapped in the present study and where *SPP1* lies.

6.1.3. Secreted phosphoprotein 1 or SPP1

As described in Chapter 1, SPP1 is a component of the ECM that interacts with cell surface receptors, including integrins, to mediate cell adhesion, migration, differentiation, survival and immune function (Garlow *et al.*, 2002). Integrins are adhesion molecules that have been implicated in the porcine implantation cascade, and SPP1 has been found at the conceptus-maternal interface prior to implantation (Garlow *et al.*, 2002). *SPP1* is up regulated in the uterus during early pregnancy in

humans, mice, rabbits, goats, sheep and pigs (Johnson *et al.*, 2003a), and it is a key component of implantation. Thus, the regulation and function of SPP1 has been shown to be temporally and spatially implicated in the establishment and maintenance of a successful pregnancy (Johnson *et al.*, 2003a).

Multiple integrin receptors for SPP1 are present on trophoblast and endometrial LE of humans and domestic animals, some of which increase during the peri-implantation period (Burghardt *et al.*, 2002; Johnson *et al.*, 2003a; Kim *et al.*, 2010). Thus, SPP1 has been found to bind the αvβ6 and possibly αvβ3 integrins expressed by trophoblast and uterus to induce focal adhesion assembly as a prerequisite for adhesion and migration of trophoblast cells in different species (Johnson *et al.*, 1999a; Apparao *et al.*, 2001; von Wolff *et al.*, 2001; Apparao *et al.*, 2003; White *et al.*, 2006; Kim *et al.*, 2010).

All this together forms substantial evidence to support the function of *SPP1* in implantation and placentation of mammalian species. However, the temporal and spatial pattern of SPP1 mRNA and protein expression and regulation is complex and species specific.

6.1.3.1. Pig expression of SPP1

In the pig, the conceptus secretes oestrogen beginning on days 11 and 12 for pregnancy recognition and to prevent luteolysis, as well as to activate a number of endometrial growth factors and cytokine mediators of conceptus attachment and implantation (White *et al.*, 2005). This conceptus oestrogen induced SPP1 in LE, directly adjacent to the conceptus, where it potentially binds αvβ3 integrin at the apical surface of LE and αvβ6 on trophoblast to mediate attachment for implantation in pigs (White *et al.*, 2005; Erikson *et al.*, 2009; Ka *et al.*, 2009). This suggests a direct conceptus influence that is not evident in other mammals (Garlow *et al.*, 2002).

Previous work has defined the temporal and spatial expression of SPP1 mRNA and protein during pregnancy in the pig uterus, and its interaction with integrins,

coinciding with important events, such as conceptus elongation, implantation and placentation, indicating an important function or role of SPP1 in this process (Garlow *et al.*, 2002). *SPP1* mRNA was first evident in endometrial LE between day 12 and 15 of pregnancy, and along the entire LE thereafter, with a 20-fold increase between day 25 and 85 of pregnancy (Garlow *et al.*, 2002). However, *SPP1* mRNA is not present in the glandular epithelium (GE) until between days 30 and 35 of pregnancy (Garlow *et al.*, 2002; White *et al.*, 2005), and increases between days 40 and 85. This expression in the GE appears to be regulated by placental progesterone production (White *et al.*, 2005). *SPP1* mRNA was not detected in porcine trophoblast. After day 30 of pregnancy, expression is then maintained in both LE and GE throughout gestation, resulting in SPP1 protein along the entire uterine-placental interface (Garlow *et al.*, 2002; Burghardt *et al.*, 2002; White *et al.*, 2005).

The fact that the endometrial LE remains intact throughout pregnancy in pigs, and that there is a constant supply of SPP1 at the conceptus-maternal interface for the duration of pregnancy make the pig an excellent model to study SPP1 (Garlow *et al.*, 2002; White *et al.*, 2005; Erikson *et al.*, 2009; Bailey *et al.*, 2010).

6.1.3.2. SPP1 in other species

SPP1 appears to also play a key role in conceptus implantation and maintenance of pregnancy in sheep (Johnson *et al.*, 1999a). Expression of *SPP1* mRNA increases in the uterine glands of pregnant ewes beginning at day 13, and protein is present in uterine flushings from day 15 of pregnancy (Johnson *et al.*, 1999a). In sheep, in contrast with pigs, *SPP1* is not present in the LE, but only in the GE (Johnson *et al.*, 1999b)

In humans, *SPP1* expression is restricted to the secretory phase glands of both non-pregnant and pregnant human endometrium, and an increase in the expression of SPP1 and αvβ3 and α4β1 integrins was detected in LE during the window of implantation, around 9 days after ovulation (Apparao *et al.*, 2001; Kao *et al.*, 2002). In humans SPP1 mRNA and protein expression is restricted to the GE during the secretory phase, as in sheep.

In rabbits, *SPP1* mRNA is expressed in a stage-specific manner in the endometrium, increasing during the peri-implantation period (Apparao *et al.*, 2003), whereas in mice, it is transiently induced by oestrogen in the endometrial LE of pregnant individuals during the attachment phase of implantation (Nomura *et al.*, 1988), and SPP1 protein is prominent at the apical LE surface (White *et al.*, 2006).

6.1.4. Characterisation

For the characterisation of SPP1 in reproductive tissues, three different techniques were chosen: RT-qPCR for quantification of mRNA expression, IHC and ISH for location and quantification of protein and mRNA in the tissue, respectively. In the previous Chapter (Chapter 5), the techniques used for the collection of the tissues, as well as for the mRNA isolation and for the normalisation process for the mRNA quantification were described in detail.

Previous studies of SPP1 in these tissues were examined for the methodology used for both IHC and ISH (Garlow *et al.*, 2002; Johnson *et al.*, 2003c; White *et al.*, 2005). The protocol for the IHC, a very efficient technique, was adapted from these previous studies, and optimised for the fixation techniques used in this study. As mentioned previously, the thrombin-cleavage of SPP1 gives rise to two fragments; a 45 KDa fragment containing the amino half terminal that is recognised by LF-124 antiserum, and a 25 KDa protein that contains the carboxyl half form that is recognised by LF-123 (or LF-166) antiserum. Both recognise the 70 KDa native protein. In the ovine uterus differences have been found in the location of these fragments (Johnson *et al.*, 2003b), and for this reason a cocktail of two antibodies was used for later studies of SPP1. As a secondary antibody, a fluorescent labelled antibody was chosen. In the present study, the antibodies for SPP1 were the same used previously, not only in pigs but also in sheep.

For the mRNA location, the technique used was also based on previous studies (White *et al.*, 2005), where results were satisfactory with frozen tissues. In this study, paraffin embedded tissues were used, where an approximate 25% mRNA signal loss

is expected. However, ^{35}S riboprobes represent one of the most sensitive methods for the detection of mRNA in tissue sections. This technique allows the precise cytological location of a nucleic acid sequences (target genes) to be determined in their cellular environment. Genes and transcripts can be localised on a section of tissue by incubating the slides with a radiolabelled RNA/DNA probe generated from a linearised plasmid with the template of interest.

Sense and antisense hybridisations are widely used for ISH with riboprobes. Messenger RNA is normally synthesised from chromosomal DNA in the 3' to 5' direction, producing sense mRNA. Thus, in order to make an antisense riboprobe the cDNA insert is sub-cloned in a transcription vector in the 3' to 5' direction relative to an RNA polymerase initiation site. The transcription takes place in the presence of a labelled nucleotide, and thus, the label allows the localisation of the mRNA in the sample.

6.2. Materials and Methods

The validation of part of the methodology used in this Chapter was described in Chapter 5.

6.2.1. Tissue samples

Reproductive tissues were collected for this study from two different populations, forming two samples sets. Tables 6.1 and 6.2 summarise relevant information for these populations. Further information on these populations is in Appendix 4. Table 6.3 contains the mean and standard error (SEM) for some of the information that characterise these populations.

6.2.1.1. Tissue collection from Large White x Landrace crossbred gilts/sows at The Roslin Institute

The first samples were collected from nine pigs from the Roslin Institute population, a LW-LR crossbred. These pigs were a mix of gilts and sows in their first and second parity, and the tissues were collected from 41 to 46 days of pregnancy (5.2.1).

6.2.1.2. Tissue collection from Large White and Meishan gilts at INRA (France)

Reproductive tissues from four MS and five LW pregnant gilts were collected at the INRA (Institut national de la recherché agronomique) Experimental Station "Le Magneraud" (Surgère, Charente-Maritime) in France. The collection of tissues was done as previously described in Chapter 5 (5.2.1) with some minor differences, as described below.

All the animals were artificially inseminated with semen from purebred boars from the corresponding breed after their oestrous cycles were synchronised with Regumate (0.04% altrenogest) (Janssen, Issy Les Moulineaux, France) on their food for 18 days. On days 41-42 of pregnancy, the animals were slaughtered by electric shock, and once death was confirmed, they were suspended by the hind limbs and a mid-ventral incision through the skin, fat and body wall was performed.

The pregnant reproductive tract was lifted out of the body cavity and removed by cutting through the vagina, ensuring that both ovaries were retained. The tract was collected in a dissecting tray and transferred to the dissecting area. The ovaries were transferred to a pot containing physiological saline for later dissection and counting of CL. The collection of tissues from the smallest and a normal sized foetus in the litter were performed as previously described in Chapter 5 (5.2.1).

Pig ID	574	509	Y24	W12	Y22	W2	W8	Y26	W7
Parity	1	1	2	1	1	2	2	2	2
Age at slaught (days)	256	354	399	397	399	403	400	396	400
Stage of pregnancy (days)	42	41	45	46	44	44	43	41	42
Litter size	15	12	16	16	13	18	17	10	12

Table 6.1 Summary of information from gilts/sows slaughtered at The Roslin Institute.

Pig ID	94879	94897	94953	94956	95022	95531	95535	95537	95580
Breed	LW	LW	LW	LW	LW	MS	MS	MS	MS
Age at slaught (days)	297	296	295	295	294	240	240	240	225
Ovulation Rate	21	29	22	22	18	23	23	23	16
Stage of pregnancy (days)	42	42	42	42	42	41	41	42	41
Litter size	12	9	20	13	16	17	17	15	15

Table 6.2 Summary of information from gilts slaughtered at France in an INRA farm. LW, Large White, MS, Meishan.

	Large White- Landrace			Large White			Meishan		
	Range	**Mean**	**± SEM**	**Range**	**Mean**	**± SEM**	**Range**	**Mean**	**± SEM**
Age of gilt (days)	103 – 256	378.22	16.08	294 – 297	295.4	0.50	225 – 240	236.25	3.35
Weight average foetus (g)	12 – 23.7	16.70	1.34	11.57 – 15.78	13.07	0.72	10.32 – 10.92	10.70	0.14
Weight smallest foetuses (g)	6.09 – 19.58	13.03	1.45	9.40 – 12.65	10.89	0.55	8.35 – 8.9	8.51	0.13
Mean litter weight (MLW)	11.31 – 22.86	15.85	1.39	12.40 – 14.57	13.13	0.38	10.18 – 11.09	10.44	0.21
SD MLW/within-litter variation	1.04 – 3.01	1.68	0.20	0.68 – 1.34	1.06	0.11	0.65 – 1.00	0.85	0.08
Weight smallest as % of MLW	53.83 – 91.59	81.01	3.72	75.14 – 95.47	82.91	3.79	75.24 – 87.39	81.61	2.48
Litter size	10 - 18	14.33	0.90	9 - 20	14.00	1.87	15 – 17	16.00	0.58
Stage of pregnancy	41 - 46	43.00	0.62	42	42.00	0.00	41 – 42	41.25	0.25
Ovulation rate	-	-	-	18 - 29	22.40	1.81	16 – 23	21.25	1.75

Table 6.3 Summary, by population, of the gilts/sows and the foetuses collected. The Table indicates the range of values, the mean, and the standard error of the mean (± SEM) separately for each measure for the three breeds or crosses used in this study.

In contrast to the sample from Roslin, this sample was more uniform; all the pigs were gilts, i.e., the tissues were collected from foetuses of the first parity. The age of the gilts was very similar and the stage of pregnancy differed in only one day.

6.2.2. Tissue preservation and processing

The tissue preservation methods used were described in Chapter 5 (5.2.2). Tissues collected in France were preserved in liquid nitrogen, RNA*later* and Methacarn, as described in Chapter 5.

6.2.3. RNA isolation, spectrophotometry and RNA quality control

The RNA isolation for the different reproductive tissues from the nine LW-LR crossbred, five LW and four MS gilts/sows was performed as described in Chapter 5 (5.2.9). The tissues collected and used for the RNA isolation for the first two LW-LR gilts/sows were preserved in liquid nitrogen. The remaining tissues used for the RNA isolation were preserved in RNA*later*.

The RNA concentration and quality was assessed as described in Chapter 5 with the Nanodrop (5.2.3) and the Agilent Bioanalyser (5.2.4), respectively. The criteria for acceptance of a sample were also defined in Chapter 5. If a tissue failed to pass the criteria after several repeats, the RNA isolation was repeated using tissue preserved by snap freezing in liquid nitrogen.

6.2.4. Reverse transcription (RT)

The RNA isolated was used as template to synthesise cDNA for the expression study. An amount of 1.25 μg of RNA was used for this synthesis. The reaction was performed as described in Chapter 5 (5.2.12).

6.2.5. Reverse Transcription-quantitative PCR (RT-qPCR)

RT-qPCR analysis was performed in the three different tissues collected for the two sets of samples separately. The cDNA from the France samples was run in duplicate and the qPCRs were performed for both duplicates. The efficiency of the amplification reaction was verified for each gene (Table 6.4) with values around 2, indicating the fine quality of the amplification.

Amplification efficiency		
Genes	**Roslin**	**France**
RPL4	1.97	1.95
HPRT1	1.92	2.02
GAPDH	1.97	2.01
SPP1	1.94	1.97

Table 6.4 Amplification efficiencies for each gene for both sets of samples. These values were calculated as described in Chapter 5 and were used for the transformation of the data.

The cDNA samples prepared from the RNA were used in a qPCR reaction in order to quantify the transcription of the gene of interest and the internal control genes. Quantitative-PCRs were run for *SPP1*, as gene of interest and for *GAPDH*, *RPL4* and *HPRT1* as internal control genes for all the samples. The selection of these control genes was described in Chapter 5 (5.2.18). In order to generate a standard curve, standards were prepared as described in 5.2.15, pooling the cDNAs for all the samples together and preparing serial dilutions of the pool (1:4, 1:8, 1:16, 1:32, 1:64, and 1:128). The cDNA of each sample was diluted 1:20.

The tissues from the nine gilts/sows from The Roslin Institute were analysed together in a sample maximisation setup, where a control per plate was included, together with the dilutions for the standard curve, the samples and a control per pig for cDNA. The samples from the tissues collected in France were run in the same way in a separate experiment.

As indicated previously, the results were verified for possible differences between duplicates, any samples with errors and to verify that the controls were negative. If any error was found the analysis for the sample with ambiguous results was repeated. Once all the samples had an appropriate result for the qPCR, the data were exported from the Mxpro software and the transformations required for the calculation of a normalisation factor for *SPP1* were performed in an Excel worksheet as described in 5.2.18.

Briefly, the Ct duplicates for each sample for each gene were averaged, and the standard deviations (SDCt) calculated. The mean Cts were transformed to relative quantities (Q) using the amplification efficiency in the following formula described previously in Chapter 5 (5.2.18).

$$\mathrm{Q} = \mathrm{E}^{(\mathrm{minCt}-\mathrm{sampleCt})} = \mathrm{E}^{\mathrm{deltaCt}}$$

A normalisation factor was calculated with the Q of the internal control genes. As shown in the following formula, the geometric mean of the Qs of these genes was calculated for each sample. The use of the geometric mean allowed for the control of possible outlier values and abundance differences between the different genes.

$$NFn = \sqrt[n]{\mathrm{REF1} + \mathrm{REF2} + \mathrm{REF3}}$$

These normalisation factors were rescaled, dividing each of the samples normalisation factors by the geometric mean of all the normalisation factors for all the samples, in order to distribute these values around unity. With these values the normalised expression level for the gene of interest, *SPP1*, was calculated as indicated in the following formula.

$$\text{Normalised } SPP1 = \frac{\text{Raw } SPP1 \text{ Q}}{\text{Normalisation Factor}}$$

The normalised *SPP1* was used in a statistical analysis for the expression level assessment.

6.2.6. Sectioning

Sections from both sets of tissues were obtained as described in Chapter 5 (5.2.16). In this case, the sections for the IHC experiment were prepared in polysine or Superfrost plus slides (Thermo Scientific). A total of 6 slides of sections were obtained, from which initially three (two as a replicate measurement and one as control) were used, and the rest were stored for possible repeats. In contrast, for the *in situ* analysis, a more complex methodology was used. The polysine slides (Thermo Scientific) were treated with H_2O containing 0.1% DEPC (Diethylpyrocarbonate, Sigma-Aldrich) for 15-30 minutes to destroy any RNase attached to them. The DEPC-H_2O was removed from the dish. The rack and the dish were wrapped in aluminium foil and baked in an oven at 180°C for between 2 hours to 4 hours, and then cooled to room temperature. The area where the tissues were sectioned, including the microtome and the water bath, was cleaned with RNase*Zap*, and the water for the water bath was DEPC-treated water. The tissue sections were placed in the treated slides and the section were placed in the oven at 60°C overnight. A total of 8 slides were obtained, from which 4 were initially used (2 antisense, one sense and one control), and the others were stored.

6.2.7. Immunohistochemistry

Immunohistochemistry studies were performed in order to detect and locate SPP1 protein on paraffin-embedded tissue sections. This required the incubation of the slides with an antibody specific to SPP1, followed by an Alexa-Fluor ® secondary antibody (Invitrogen) incubation and a counterstaining with Propidium iodide (Sigma-Aldrich, 10ng/ml).

The slides were prepared by incubating as follows: three times for 2 minutes in Xylene (Fisher), 2 minutes in 100% ethanol (Fisher), 2 minutes in 80% ethanol, 2 minutes in 50% ethanol, 5 minutes in dH_2O and twice for 5 minutes in 0.3% Tween-20 in PBS (Phosphate-buffered saline, Oxoid, Hampshire, UK). The last step was to permeabilise the tissues.

In order to reduce background, non-specific antibody binding sites were blocked by incubating the slides for 1 hour at room temperature in antibody dilution buffer (2 parts 0.02M PBS, 1.0% BSA (Bovine Serum Albumin, Sigma-Aldrich), 0.3% Tween-20 and 1 part of Glycerol) containing 10% normal goat serum (Sigma-Aldrich). After this incubation, the slides were rinsed in PBS for 5 minutes.

A cocktail of rabbit anti-human recombinant SPP1 serum (LF-166 and LF-124) (Fisher *et al.*, 1995) diluted in antibody dilution buffer containing 10% normal goat serum (Vector Labs, Peterborough, UK) was used for the primary antibody incubation. As negative control, Normal Rabbit IgG (Insight Biotechnology Limited, Middlesex, UK) was used, also prepared in antibody dilution buffer containing 10% normal goat serum. Each slide was treated with 50 µl of this mix. The slides were covered with a parafilm cover slip, and incubated in a humidified chamber overnight at 4°C or 1 hour at 37°C. After incubation the slides were rinsed three times in PBS for 5 minutes each time.

As for the primary antibody, the secondary antibody (Alexa Fluor 488 ® goat anti-rabbit IgG, Invitrogen) was prepared in antibody dilution buffer containing 10% normal goat serum. The slides were placed in a light proof box and away from the sun light and 50 µl of the secondary antibody mix was applied to each slide. The slides were incubated for 1 hour at room temperature in the dark. After the incubation the slides were rinsed three times in PBS for 5 minutes each.

In order to recognise cells, the nuclei/nucleus area of the cells was counterstained with Propidium iodide (PI) which fluoresces red under a green light. The slides were placed in PI for 2-3 minutes and washed twice for 1 minute in PBS and twice for 1 minute in dH_2O.

The excess moisture was removed from the slides, which were then laid on a dry filter paper in a light proof box. A drop of ProLong ® Gold Antifade reagent (Invitrogen), warmed to room temperature before use, was added to the specimens and they were covered with cover slips. The samples were left on a flat, dry surface to cure for at least 24 hours at room temperature in the dark. After the cure, the edges

of the cover slip were sealed with nail polish to retard the oxidation and extend the life of the samples, and the slides were stored at room temperature. The results were observed and analysed with a Nikon Inverted Laser EC-1 confocal microscope.

6.2.8. Agarose gel

The different concentrations of agarose gel used for the preparation of the probe for the ISH were prepared and run as described in 5.2.14.

6.2.9. *In situ* Hybridisation

A plasmid vector containing a piece of DNA from the *SPP1* gene was built to produce an antisense and a sense template, which were used to generate probes. These probes were labelled with ^{35}S-UTP and used in an ISH study to locate *SPP1*.

6.2.9.1. Development of probe templates for in situ *Hybridisation*

SPP1 *cDNA preparation and purification*

Three RNA samples were chosen, the cDNA was prepared and *SPP1* was amplified by PCR reaction. The size of the amplified fragment was verified in a 3% agarose gel with the help of a ladder (1,000 bps), and the bands of the correct size (120 bps) were excised from the gel. The DNA was purified from these bands with a QIAquick Gel extraction kit (QIAGEN).

Briefly, the gel slices were weighed and three volumes of buffer QG were added to one volume of gel. The mix was incubated at 50ºC for 10 minutes to dissolve the gel, vortexing a couple of times during the incubation. Once the gel was dissolved, a gel volume of isopropanol was added to the sample and mixed. A QIAquick spin column was placed in a 2 ml collection tube and a maximum of 800 µl of the sample was applied to the column where the DNA was bound, and the column was centrifuged at 17,900 x *g* for 1 minute. This step was repeated with the same 800 µl of sample, and then the flow-through was discarded. The rest of the sample was applied to the

column and the previous steps were repeated. The column was washed with 0.75 ml of Buffer PE and centrifuged for 1 minute. The flow-through was discarded and the column was centrifuged for an additional minute at 17,900 x *g*. The column was placed in a clean 1.5 ml microcentrifuge tube, and 50 µl of water were added to the centre of the membrane to elute DNA. The tube was centrifuged for 1 minute and the 50 µl applied to the column again and the tube was centrifuged. This step was repeated with the same 50 µl of water. The DNA concentration was measured in the Nanodrop (5.2.10).

Ligation into pGEM-T easy

The DNA fragment corresponding to the *SPP1* cDNA was inserted in a vector in a ligation reaction. The pGEM-T Easy Vector (Figures 6.1 and 6.2) (Promega Corporation) was used in this ligation following the protocol of the vector. In order to make sure that the ligation was efficient, three different insert:vector molar ratios were prepared for the three samples and both positive and background controls were also prepared. The quantity (ng) of insert needed for each ratio was calculated (Table 6.5) with the following formula for the optimisation of insert:vector molar ratio.

$$\frac{\text{ng of vector x kb size of insert}}{\text{kb size of vector}} \text{ x insert: vector molar ratio} = \text{ng of insert}$$

where the concentration of the vector was 50 ng and the kbp size of the insert was 120 bp and the size of the vector was 3,015 bp. The insert: vector molar ratios used were 3:1, 1:1 and 1:3.

Ligation	Insert		
	3:1	**1:1**	**1:3**
Sample 1	5.97 ng	1.99 ng	0.66 ng
Sample 2	5.97 ng	1.99 ng	0.66 ng
Sample 3	5.97 ng	1.99 ng	0.66 ng
Positive control	Control insert DNA (2 µl)		
Background control	-	-	-

Table 6.5 Summary of samples and controls in the ligation reaction. The Table indicates the ng needed of insert for each insert:vector ratio reaction, and the controls (positive + background) included in the reaction.

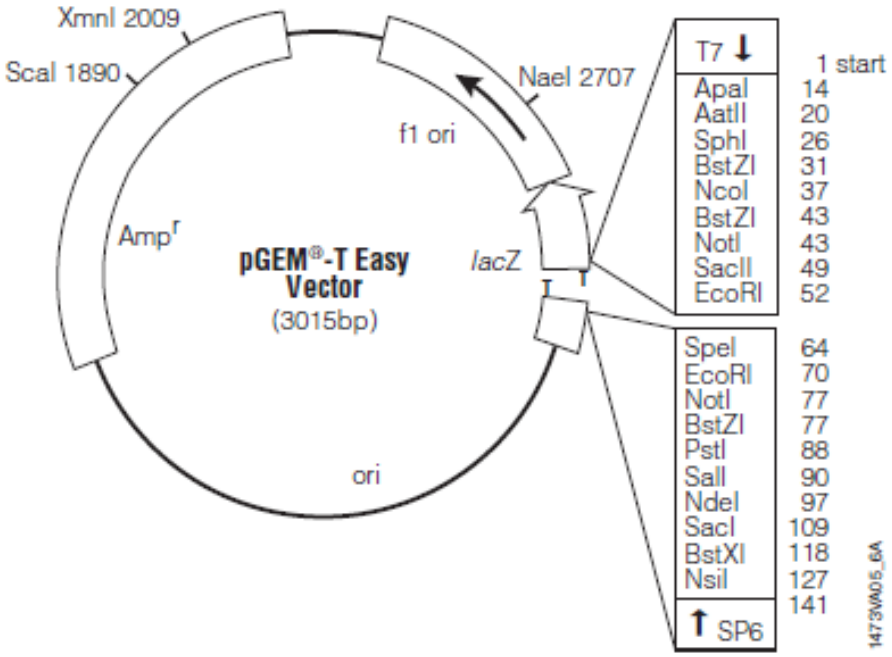

Figure 6.1 pGEM-T Easy Vector map. This Vector contains a T7 and SP6 RNA polymerase transcription initiation sites and promoters, a multiple cloning region or open reading frame, a LacZ start codon, a Lac operon, β-lactamese coding region, lac operon sequences, and an ampicillin antibiotic-resistant gene. The LacZ gene has the cloning region, meaning that the successful ligation of an insert into the vector interrupts the coding sequence of β-galactosidase resulting in white colonies in an X-gal plate. Figure taken from Promega pGEM-T easy vector protocol.

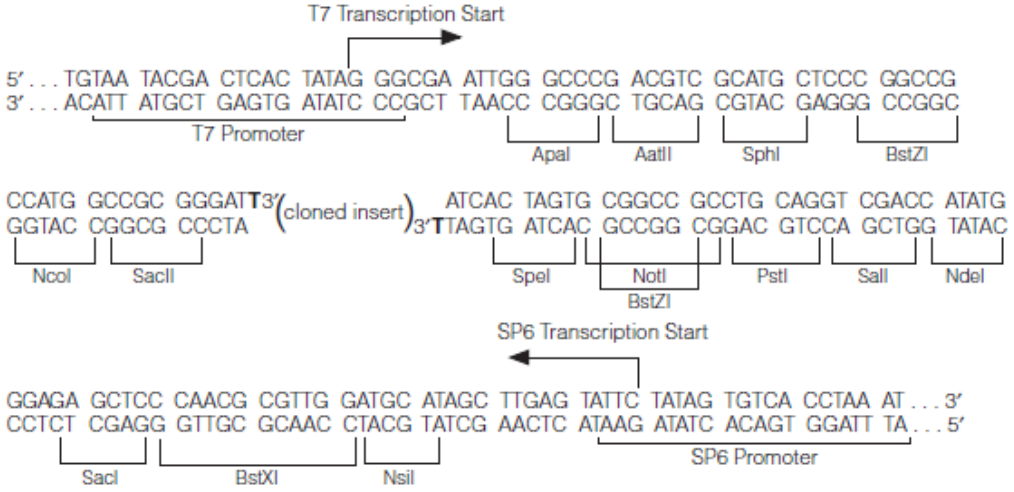

Figure 6.2 The promoter and multiple cloning sequence of the pGEM-T Vector (Promega Corporation). The top strand corresponds to the RNA synthesised by T7 RNA polymerase. The bottom strand corresponds to the RNA synthesised by SP6 RNA polymerase. The restriction enzymes and the restriction sites are indicated.

The reactions were performed in 0.5 ml tubes where 5 µl of 2x rapid ligation buffer was added to the tube, followed by 1 µl of the vector and the correspondent amount of insert DNA for the standard reaction, 2 µl of control insert DNA (from the Kit) for the positive control (positive control, no PCR product), and nothing for the background control. In a last step, 1 µl of T4 DNA ligase (kit) was added to the mix. The reactions were mixed by pipetting and the tubes were incubated overnight at 4ºC.

Transformation with JM109 competent cells

The vector with the insert was used to transform JM109 high efficiency competent cells (Promega Corporation). After the incubation, the ligation reactions were centrifuged and 2 µl of each ligation reaction was added to a sterile 1.5 ml tube on ice. Meantime, the JM109 competent cells were placed on ice until just thawed and mixed gently by flicking the tube. A 50 µl aliquot of the cells was transferred to the ligation reaction tubes. Another tube was prepared with 0.1 ng of uncut plasmid and a 100 µl aliquot of cells for determination of the transformation efficiency of the competent cells. The tubes were gently flicked and incubated on ice for 20 minutes.

The cells were heat-shocked for 45-50 seconds in a water bath at exactly 42ºC, in order to allow the vector to enter the cells, and the tubes were returned to ice for 2 minutes. Room temperature lysogeny broth (LB) was added to the ligation reaction transformations and to the uncut DNA control tube, 950 µl and 900 µl, respectively. The tubes were incubated for 1.5 hours at 37ºC with shaking (~150rpm). In order to see the results of the transformation, 100µl of each culture was plated into duplicate LB plates. For the uncut plasmid control, a 1:10 dilution prepared with SOC (super optimal broth) was plated. These plates contained X-Gal (Promega Corporation), ampicillin (Sigma-Aldrich) antibiotic and IPTG (isopropyl-beta-D-thiogalactopyranoside) (Promega Corporation). IPTG is a highly stable synthetic analogue of lactose. It inactivates the lac repressor and induces synthesis of the lac operon. The X-Gal was used to verify if the *LacZ* gene, which expresses β-galactosidase, was interrupted by the insert. X-Gal was cleavage by β-galactosidase and one of the products of this reaction was oxidised giving the blue colour to the colonies. The plates were incubated overnight at 37ºC.

Selection of colonies

After the incubation, white colonies, which contain the insert, were selected. Thanks to the X-Gal, non-transformed colonies grew to be blue because the *LacZ* gene was not disrupted and transformed colonies, with the insert interrupting the *LacZ* gene, grew to be white (Table 6.6).

Two colonies were picked up from each of the three plates for each tissue, and one colony from each control, i.e. positive, background and transformation controls. These colonies were grown in 10 ml of liquid LB with ampicillin overnight at 37 ºC with shaking.

Ligation	No. of colonies
Sample 1	2
Sample 2	2
Sample 3	2
Positive control	1
Background control	1
Transformation control	1

Table 6.6 Summary of the colonies collected per sample and control

Minipreps

A Wizard® plus SV minipreps DNA purification System (Promega Corporation) kit was used to isolate the recombinant plasmid DNA. This kit contained different solutions (Cell resuspension solution, cell lysis solution, Alkaline protease solution, Neutralisation solution, and Column wash solution), spin columns and collection tubes. After the incubation, two 1.5 ml microcentrifuge tubes were filled with each cell culture and centrifuged for 5 minutes at 10,000 x *g*. The supernatant was poured off, and the tube was inverted on a paper towel to remove any excess media. Cell Resuspension Solution (250 µl) was added to one of the tubes for each transformation and to the control tubes. The pellet was completely resuspended by pipetting and the resuspension was transferred to the next pellet of the same transformation. Cell Lysis Solution (250 µl) was added and the suspension was mixed by inverting the tube four times, since vortexing from this step onward could provoke shearing of chromosomal DNA. The suspension was incubated for between 1 and 5 minutes until the cell suspension cleared. Alkaline Protease Solution (10 µl) was added and mixed by inverting the tubes four times. In order to lyse the cells, the mix was incubated for a maximum of 5 minutes at room temperature. Neutralisation Solution (350 µl) was added and mixed by inverting the tube four times. The cell lysate was centrifuged at maximum speed (around 14,000 x *g*) for 10 minutes at room temperature. This centrifugation resulted in a clear lysate and a white pellet.

The clear lysate was decanted (approximately 850 µl) carefully into a Spin Column placed into a 2 ml collection tube, one for each sample. The column was centrifuged at maximum speed for 1 minute at room temperature and the flow through was

discarded. Column Wash Solution (750 µl) was added to the Spin column and the column was centrifuged at maximum speed for 1 minute at room temperature. The flow through was discarded and the previous wash was repeated with 250 µl of Column Wash Solution. The column was centrifuged at maximum speed for 2 minutes at room temperature, the flow through was discarded. The column was transferred to a new, sterile 1.5 ml microcentrifuge tube and centrifuged for another minute at maximum speed to make sure no Column Wash Solution was transferred. Once again, the column was transferred to a new tube. In order to elute the plasmid DNA, 100 µl of nuclease-free water were added to the Spin column and the column was centrifuged at maximum speed for 1 minute at room temperature. The column was removed from the tube and discarded. DNA concentration was measured with the Nanodrop and the sample was stored at -20ºC.

Digestion with restriction enzyme

In order to make sure the required insert was ligated into the vector, the size of the fragment inserted into the plasmid needed verification. An enzyme with restriction sites at both ends of the insert was chosen to cut the plasmid in two fragments, one for the plasmid and one for the insert. The enzyme used was NotI (BioLabs, New England) and the restriction sites are shown in Figure 6.3. All the DNA samples extracted from the transformation were digested.

Figure 6.3 NotI restriction size. One unit is defined as the amount of enzyme required to digest 1µg of pBC4 DNA in 1 hour at 37ºC in a total reaction volume of 50 µl.

For the digestion reaction 2 µl of 10x NE buffer 3 (100 mM NaCl, 50mM Tris-HCl, 10 mM MgCl2, 1 mM DTT, pH 7.9 at 25ºC), 2 µl of BSA (1mg/ml), 15 µl of Plasmid DNA (0.2-1 mg/ml) and 10 units of NotI (1 µl) were mixed and incubated at 37ºC for 60 minutes. From this digest, an amount of 10 µl was run on a gel to verify the size of the resulting fragments. The size of the desirable insert, *SPP1* gene

cDNA, was 120 bp and the pGEM-T Easy Vector consisted of 3,015 bp. The insert together with a part of the cloning regions after cutting with NotI was a total of 154 bp and the rest of the vector was 2,981 bp. Each digest was run next to 7.5 µl of the correspondent uncut plasmid-insert ligation in a 2% agarose gel with a Hyperladder I (up to 10,000bp) (Bioline, London, UK) as a size marker. The gel was inspected for bands matching the size of the insert fragment and the plasmid fragment. The samples where these two fragment were found, were re-run in a 3% agarose gel with a PCR marker (Promega Corporation) (1,000 bp), in order to confirm the size of the small fragment, the *SPP1* insert. Once the size was confirmed, the DNA for the vector with the right size insert was sent (DNA Sequencing and Services, Dundee, UK) to be sequenced in order to verify that it was the *SPP1* cDNA (Accession number NM_214023). An aliquot of the culture from which the isolated plasmid DNA resulted in the right fragments size was preserved in 50% glycerol for long term storage.

Minipreps to isolate template

Once the sequence was confirmed, the cells, kept in glycerol, were plated in two LB plates containing only Ampicillin and incubated overnight at 37ºC. Next day, one of the colonies from the plates was picked and grown in 25 ml of liquid LB with ampicillin overnight at 37ºC with shaking (250 rpm), in order to isolate a large amount of plasmid DNA by minipreps.

After the incubation, the liquid LB was distributed in 1.5 ml microcentrifuge tubes and centrifuged for 5 minutes at 10,000 x *g*. The supernatant was poured off and the tubes were inverted on a paper towel to remove any excess media. Cell Resuspension Solution (Promega Corporation) (250 µl) was added to half of the tubes and the pellet was completely resuspended by pipetting. The resuspension was transferred to the other half of the pellets.

Cell Lysis Solution (250 µl) was added and the solution was mixed by inverting the tube four times as previously described. The suspension was incubated for 1 minute to a maximum of 5 minutes until the cell suspension cleared. Alkaline Protease

Solution (10 µl) was added and mixed by inverting the tubes four times. The mix was incubated for a maximum of 5 minutes at room temperature in order to lyse the cells. Neutralisation Solution (350 µl) was added and mixed by inverting the tube four times. The cell lysate was centrifuged at maximum speed (around 14,000 x *g*) for 10 minutes at room temperature.

The cleared lysate was decanted (approximately 850 µl) carefully into a Spin Column placed into a 2ml collection tube. The column was centrifuged at maximum speed for 1 minute at room temperature and the flow-through was discarded. Column Wash Solution (750 µl) was added to the Spin Column and the column was centrifuged at maximum speed for 1 minute at room temperature. The flow-through was discarded and the previous wash was repeated with 250 µl of Column Wash Solution. The column was centrifuged at maximum speed for 2 minutes at room temperature, the flow-through was discarded and the column was centrifuged for another minute at maximum speed, to make sure there was not any Column Wash Solution remaining. After this step, the column was transferred to a new, sterile 1.5 ml microcentrifuge tube. In order to elute the plasmid DNA, 100 µl of nuclease-free water were added to the Spin Column and the column was centrifuged at maximum speed for 1 minute at room temperature. The column was removed from the tube and discarded. The resulting DNA concentration was measured with the Nanodrop and the sample was stored at -20ºC.

Preparation of antisense and sense templates

The plasmid sequence was examined in order to verify the orientation of the inserted DNA and two digestions were set up: one using a restriction enzyme in the poly-cloning region on one side of the insert and one on the other side of the insert. The restriction enzymes were unique in the plasmid and did not cut the insert itself. This means that with each digestion a unique cut was done getting the two different templates (i.e., the antisense and the sense) needed for the preparation of the probe.

A riboprobe in vitro transcription system was used for the preparation of the probes, and one of the requirement was that the enzyme used for the cut did not leave a 3'

overhang. This is required in order to avoid the appearance of extraneous transcripts that can contain sequences complementary to the expected transcript as well as sequences corresponding to vector DNA. Thus, enzymes of this kind or enzymes with recognition sequences at both sides of the insert were not adequate, and this included EcoRI, NotI, and BstZI.

The digestion with NcoI (Roche) enzyme gave rise to the antisense probe which was synthesised with SP6 RNA polymerase, generating the complementary strand to the mRNA. SalI (Roche) enzyme gave rise to the sense probe, which sequence was equal to the mRNA and synthesised with T7 RNA polymerase. An aliquot of 75 µl of plasmid was cut with 40 units of NcoI for the antisense or SalI for the sense (4 µl of restriction enzyme (10 units/µl)) and 10 µl of 10x digestion buffer of each enzyme, incubating for 1 hour at 37ºC and then overnight at room temperature. A 5 µl aliquot of the cut plasmid was initially run on a 1.2% Agarose gel. If any extra bands were discovered as a result of uncut plasmid, the entire digestion was run in a gel, the bands were excised from the gel and the plasmid was extracted using the QIAquick gel extraction Kit as previously described in the cDNA purification step.

After the extraction, the resulting eluate was mixed with RNase free water up to 80 µl if needed or a multiple of 80 for the next step, a DNA extraction with a PCI;CI;EtOH precipitation method. The DNA eluted from the gel was mixed with 10 µl of 10x proteinase K buffer (0.01M Tris pH 7.8, 0.05M EDTA, and 5% SDS) and 10 µl 1mg/ml Proteinase K, for a Proteinase K digestion. This mix was incubated at 65ºC for 60 minutes. The digest was extracted with an equal volume of PCI (Phenol; Chloroform; Isoamylalcohol) (Sigma-Aldrich), mixed and centrifuged for one minute. The resulting upper phase was transferred to a new tube and an equal volume of CI (Chloroform; Isoamylalcohol) (Sigma-Aldrich) was added; the tube was mixed and centrifuged for 1 minute. The resulting upper phase was transferred to a new tube and mixed with 0.1 volumes of 3M NaAc and 3 volumes of ice-cold EtOH. The tube was placed at -20ºC for 30 minutes. After the incubation, the mix was centrifuged for 15 minutes at 12,000 x *g* and 4ºC.

The pellet was washed with 70% ethanol prepared with DEPC-treated water and the tube was centrifuged for 5 minutes at full speed. The pellet was air-dried for 20 minutes and then redissolved in 10 µl of DEPC-treated water (1 µl per µg starting material). The concentration was measured using the Nanodrop and the sample was stored at -20 ºC in 2 µl aliquots. One of the aliquots was re-run in a gel in order to verify that the plasmid was the right size. The result of this step was two linearised DNA templates, one for the antisense probe and another for the sense.

6.2.9.2. Labelling of the probe with ^{35}S-UTP

In this step, the previously prepared templates were used to prepare riboprobes labelled with ^{35}S. For this, a reaction was performed where an RNA polymerase was bound to the transcription start site of a linearised DNA template and run off transcripts complementary to the template. The template was then removed by digestion with DNase and the unincorporated nucleotides were removed by passing through a spin column resulting in a probe.

For the preparation of the probe the Riboprobe Combination System Kit (Promega Corporation) containing T7 and SP6 polymerase was used. This kit also included a 5 x transcription buffer, 100 mM DTT, RNasin, 10mM rATP, 10mM rCTP, 10mM rGTP, 10mM rUTPs, and RQ1-RNase-free DNase. The reagents, excluding the enzymes, were allowed to reach room temperature for around 1 hour. Three reactions were performed in 1.5 ml screw cap tubes, one for the antisense, one for the sense probe and one for the control (pTRI-GAPDH Rat antisense control template, Ambion, Paisley, UK).

In these tubes the following reagents were combined: 4 µl of 5 x transcription buffer, 2 µl of 100 mM DTT, 1 µl of RNasin (20 units), 4 µl of diluted rNTPs (2.5 mM for each rNTP except rUTP, mixing equal volume of each and RNase-free water for the final concentration), 2 µl of template DNA (0.2 – 1mg/ml) for antisense and sense, and 1 µl of the control, 4 µl of ^{35}S-UTP (1.85 MBq) (^{35}S-labelled rUTP 462.5MBq/ml (Perkin Elmer, Massachusetts, USA)), and 1 µl of RNA polymerase (SP6 for the antisense probe and T7 for the sense and the control). RNase-free water

was added up to a volume of 20 µl; the tubes were mixed and centrifuged briefly. This mix was incubated at 37ºC for 60 minutes in a water bath.

After the incubation, the tubes were removed from the water bath and 10 µl of tRNA (10mg/ml) (Sigma-Aldrich), 6 µl of 5x transcription buffer, 1 µl of RNasin, 1 µl of RQ1-RNase-free DNase and 12 µl of RNase-free water were added to the tubes. This was the standard reaction used, but the amount was adjusted to accommodate the number of slides that needed to be treated. The tubes were mixed and microfuged briefly. From this mix, 1µl was removed and retained in another tube for determination of incorporation and specific activity. The rest of the reaction was incubated at 37ºC for 15 minutes.

During the incubation, the Illustra™ ProbeQuant™ G50 micro columns (GE Healthcare, Bucks, UK) were prepared for its later use, resuspending the resin by vortexing, loosening the cap one-quarter turn and snapping off the bottom closure. The columns were placed in a collection tube (RNase-free 1.5ml tube) and centrifuged at 800 x *g* for 1 minute. The collection tubes were changed and the probes were applied to the centre of the columns and centrifuged at 800 x *g* for 2 minutes. As a result, the collection tubes contained the finished probe. In order to measure the incorporation in the finished probe, 1µl aliquots of the finished probes were placed in different minivials, two for each probe.

The 1 µl retained before the 15 minute incubation was mixed with 9 µl of RNase-free water. The tubes were mixed and microfuged, and 1 µl of each probe was spotted onto 2 different GF/D filter discs (Whatman, Kent, UK) (two for each probe). The discs were allowed to dry and inserted in minivials to calculate the total counts. From this mix, another 1 µl of each probe was added to 2 tubes (two for each probe - duplicates) containing 89 µl of water and 10 µl of tRNA (10 mg/ml). Ice cold 5% TCA (Trichloroacetic acid, Acros organics, Geel, Belgium) (500 µl) was added to each of these tubes and they were incubated on ice for 5 minutes.

Meanwhile, two 2 ml syringes (duplicates), one for each tube, were prepared for each probe with a GF/D filter disc pushed down to the bottom. These syringes were placed

in a rack over a collection vessel and the filter discs were pre-wetted with 1 ml ice cold 5% TCA allowing it to drip through.

Once the ice incubation of the probe-tRNA-TCA mix was finished, the tubes were mixed and the solution was added to the syringes and allowed to drip through leaving a radioactive flow. The filters were washed three times with 1 ml of ice cold 5% TCA and once with 3 ml of acetone (AnalaR NORMAPUR, VWR International). After all the liquid had dripped through, the discs were removed and allowed to dry. Once dried, they were transferred to minivials to estimate incorporated counts.

Scintillation fluid (Optiphase Hisafe 3 Wallac, Perkin Elmer) (2.5ml) was added to each minivial. The minivials were capped, mixed and the outside was wiped with 70% ethanol. The vials were placed on the scintillation counter (Wallac 1410 Liquid Scintillation Counter, Perkin Elmer) and counted with the appropriate program. The average counts from the duplicate measurements of the finished probe were used to calculate the amount of labelled probe to add to the slides. The total and incorporated counts were used to calculate the percentage of incorporation and the specific activity of the probe, to estimate the efficiency of the labelling reaction. All the calculations were performed as described in the Promega Riboprobe *in vitro* Transcription Systems protocol (Promega Corporation).

6.2.9.3. In situ *Hybridisation*

The visualisation and location of the nucleic acid sequence for *SPP1* in the cellular environment was performed by ISH in tissue sections. The ISH consisted of the hybridisation of a nucleic acid probe of complementary base sequence to the target. All glassware used in the experiment was baked for 4 hours at 180°C before use. All buffers and ethanol dilutions were prepared in Diethylpyrocarbonate (DEPC)-treated water, prepared previously with deionised water treated with 1 ml DEPC (Sigma-Aldrich) per litre, left overnight and autoclaved. The PBS (Oxoid) was prepared with deionised water and treated with DEPC as for the DEPC-treated water. Gloves were worn at all times, and the surfaces and equipment were treated with RNaseZ*ap*.

When Xylene, ethanol and mounting medium were used, the sections were handled in the fume hood.

The sections were pre-treated to unmask target nucleic acids with a post fixation treatment. The paraffin wax was removed washing the slides twice in Xylene for 5 minutes. Then the sections were rehydrated through a graded series of ethanol (100% (twice), 95%, 70%, 50% and 30%) for 2 minutes each. Sections were quickly washed twice in DEPC-treated water before the post-fixation in 4% PFA (Sigma-Aldrich) in PBS for 10 minutes. They were washed twice for 2 minutes in PBS and then digested with proteinase K (20 µg/ml) (Promega Corporation) in Proteinase K digestion buffer (50mM Tris (Fisher Scientific Ltd.), 5 mM EDTA (Fisher Scientific Ltd.), pH 8) for 8 minutes at 37ºC. Sections were then washed in PBS for one minute and refixed for 5 minutes in 4% PFA, rinsed twice for 5 minutes each in PBS, rinsed for 10 seconds in DEPC-treated water, dehydrated through a graded series of ethanol for 2 minutes each (30%, 50%, 70% and 90%), and three times in 100% ethanol for 5 minutes. Then the tissue sections were allowed to dry at room temperature for 30 minutes. Once they were dry, the slides were either stored at 4ºC for up to a week or hybridised with the probe.

Sections were hybridised with freshly prepared radiolabelled cRNA probe diluted in hybridisation buffer (50% deionised formamide (Ambion), 20 x SSC (Sigma-Aldrich), 1x Denhardt's solution (Sigma-Aldrich), 10% dextran sulphate (Sigma-Aldrich), 0.5 mg/ml yeast RNA (Sigma-Aldrich), 100mM dithiothretiol DTT (Melford Laboratories, Suffolk, UK)) for a 5 x 10^5 cpm concentration per 50 µl of probe-hybridisation buffer mix. Probes were denatured at 70ºC for 10 minutes and slides were placed in a humidified chamber containing 50% formamide /1xSSC grouping antisense and sense slides together. Hybridisation buffer containing the probe (50 µl) was applied to the middle of each slide, and a parafilm coverslip was placed gently on top, ensuring that the whole section was covered and there were no bubbles. The slides were hybridised overnight at 55ºC.

For the post-hybridisation washes, the parafilm coverslips were gently peeled off from the slides and these were dipped vertically 15 times in 2 x SSC at 55ºC in a

plastic beaker to wash off any unbound probe. This wash was repeated in a second beaker. The slides were placed into a rack in a dish containing 2 x SSC/50% formamide/0.1% β-mercaptoethanol (Invitrogen) for 30 minutes at 55ºC in a water bath with shaking. The slide rack was moved to a dish containing 2 x SSC and incubated for 15 minutes at 55ºC. This wash was repeated.

The slides were digested with deoxyribonuclease (DNase)-free ribonuclease (RNase; 10 μg/ml) (Sigma-Aldrich) in 2 x SSC for 60 minutes at 37ºC to remove non-specifically bound probe. Sections were then washed as follows: 2 x SSC for 30 minutes at 37ºC; 2 x SSC for 15 minutes at 37ºC; 2 x SSC/50% formamide/0.1% β-mercaptoethanol for 20 minutes at 55ºC with shaking. After these washes, a series of washes at room temperature were performed: 2 x SSC for 15 minutes; 0.1 x SSC for 15 minutes; 60% Ethanol for 1 minute; 80% Ethanol for 2 minutes; 95% Ethanol for 2 minutes; 100% Ethanol for 2 minutes. After these washes, the slides were air dried.

Liquid film emulsion autoradiography was performed using Amersham Hypercoat Emulsion LM1 (GE healthcare) in a dark room with only a safe light on. The emulsion was melted in a water bath at 42-45ºC for 10 minutes and poured into a dipping chamber inside the water bath. Each slide was slowly and smoothly dipped into the emulsion and the back of the slide was wiped with a tissue. Slides were placed vertically in a rack and left to dry overnight in a light-proof box.

Next day, the dry slides were placed in a light-proof slide box with a bag of desiccant (Silica gel), the box was sealed with parafilm, and stored at 4ºC for a week. After this period, the slides were allowed to warm at room temperature over 30-60 minutes and developed. Three dishes were prepared for the washes inside a plastic box with water, in order to keep the solutions slightly cooler than room temperature. The first dish contained the developer, Kodak D-19 (Sigma-Aldrich), the second was deionised water, and the third contained the fixer, Kodak fixer (Sigma-Aldrich). The slides were transferred to a staining rack and treated for 3 minutes, 1 minute, and 5 minutes in the dishes, respectively. To finish the developing, the slides were washed for 10 minutes under gentle running tap water and the excess emulsion was scraped from the back of the slides with a razor blade.

The sections were counterstained in Haematoxylin solution modified according to Gill II (VWR International) for 1 minute, washed for 5 minutes under running tap water, dehydrated through a graded series of ethanol (30%, 50%, 70%, 90% and 100%) for 1 minute and twice in Xylene for 5 minutes. The sections were drained onto tissue, a drop of DPX mountant (Sigma-Aldrich) was added and a cover slip was carefully placed on top. The slides were left to dry overnight and the results were evaluated by both bright field and dark field microscopy with a Zeiss Photomicroscope III (Carl Zeiss Inc., Thornwood, NY). Image J was used to quantify signal.

6.2.10. Image analysis

For the analysis of the microscope capture image from the IHC and the ISH, Image J software was used (http://imagej.nih.gov/ij/). For both experiments, whole utero-placental unit slide pictures were analysed, since the three distinctive structures from both reproductive tissues of interest are present; glandular and luminal epithelium of the endometrium and placenta. For both experiments, the replicates for each tissue were analysed together with the sense and the respective controls. Sense and controls pictures were visually analysed previously in order to verify the proper assignment of the slide and the use of the proper probe and antibody, respectively.

In the IHC pictures, the area of interest was manually selected, measured in pixels and the colour channels were split. SPP1 protein expression represented by green fluorescence was also quantified in pixels, with a threshold defined evaluating different pictures. This threshold, defined for each set, was used across all the pictures.

A similar analysis was performed in the ISH images. The area of interest was selected in the bright field image and used to quantify the number of silver grains in the corresponding dark field image with Image J Find maxima tool. The selected area was measured in Pixels and the number of nuclei was counted. The analysis was done separately for the three tissue structures represented in the whole uterus

sections and in the two sets. All the measurements were recorded in a Excel worksheet for a later analysis.

6.2.11. Statistical analysis

The number of pixels and number of grains per cell were calculated for the IHC and the ISH images, respectively. The results from both sets of samples (Roslin and France) were analysed separately. The distribution of the results for RT-qPCR, IHC and ISH was analysed in a leaf plot in SAS (SAS Inst. Inc., Cary, NC) or R (R Development Core Team, 2005) and the normality of the data graphically checked. When needed, the data were square root, $\log_{10}$ (measure + 1) or $\log_{10}$ (measure) transformed prior to statistical analysis, in order to get a normal distribution.

The correlation between the different measurements was estimated with R. An analysis of variance was performed in order to check the differences in SPP1 mRNA and protein. SPP1 levels were analysed using a general linear model procedure (PROC GLM) in SAS. Sequential models were fitted in order to find the one that best explained the data. Terms fitted included LS, uterine position, stage of pregnancy, OR (France only), parity (Roslin only), pig id, breed (France set only), foetal size (both sets) and all possible interactions were also analysed and data were blocked for gilt/sow (both sets) to account for the common maternal environment shared by the smallest and normal-sized siblings.

6.3. Results

6.3.1. Reverse transcription-quantitative PCR (RT-qPCR)

6.3.1.1. Large White – Landrace crossbred

The mean and standard error were calculated for each tissue and size. These values showed a small difference between foetal sizes in all the tissues. In endometrium tissue, the smallest foetus showed a higher expression of *SPP1* mRNA than the

normal-sized ones. However, the relationship was opposite in the placenta, where the normal-sized foetuses registered a higher expression.

The normality of the data was verified and the normalised *SPP1* was square root transformed for the following analysis. In order to test the significance of the difference, a simple model was fitted with both size and pig id as factors. In this analysis, the only significant differences between pigs were detected in the endometrium (P=0.04) (Table 6.7).

Factors	Endometrium	Placenta	Whole uterus
Model	0.050	0.904	0.339
Size	0.832	0.614	0.378
Pig ID	0.038	0.883	0.317
R-Square	0.793	0.309	0.604
CV	19.638	44.856	29.769
Mean	0.741	0.072	0.391

Table 6.7 Results of the analysis of variance for the RT-qPCR results for *SPP1* mRNA expression in the samples from LW-LR crossbred animals. The Table shows the p-values from an F-test for the model (Model), the size and pig id, and the R-square, the coefficient of variation (CV), and the mean for the tissue measures used in the analysis, the normalised *SPP1* values.

6.3.1.2. Large White and Meishan gilts

For the France sample, the results were analysed in a similar way as for the Roslin sample. However, this sample included tissues from two different breeds. The mean and the standard error of *SPP1* mRNA expression results were calculated. There is a general higher expression of *SPP1* in MS pigs, independently of the size of the foetus. Also, in most of the tissues, the mRNA expression was higher in the normal foetuses, except in the MS endometrium, where the expression of *SPP1* was higher in the smallest foetuses.

As previously, the distribution of the normalised *SPP1* was verified and the values were $\log_{10}$ transformed. In order to test the statistical significance of the observed

differences, a general linear model analysis was performed (Table 6.8). In this analysis breed and size were used as factors and the gilt was used to block for the maternal environmental effects in the breed. When analysing the data in more detail, considering the interaction in endometrium(Table 6.9), LW normal sized foetuses showed a higher expression (P=0.02) than the smallest ones. Regarding the breeds, endometrium *SPP1* mRNA expression in the smallest foetuses was higher in the MS foetuses (P=0.02) compared to the LW ones. The results of the analysis for this population, where breeds and sizes were compared, were similar in the sense of significance to the other population. Also the interaction of breed and size was examined.

	Endometrium	**Placenta**	**Whole uterus**
Model	0.028	0.223	0.668
Breed	0.017	0.224	0.692
Breed (Pig ID)	0.057	0.187	0.481
Size	0.020	0.372	0.732
Breed*Size	0.385	0.350	0.999
R-Square	0.866	0.721	0.519
CV	-18.336	-11.580	-23.588
Mean	-0.397	-2.729	-0.972

Table 6.8 Results of the analysis of variance for the RT-qPCR results for *SPP1* mRNA expression in the samples from LW and MS animals. The Table shows the p-values from an F-test for the model, breed blocked for pig id (gilt), size, and the interaction, and the R-square, the coefficient of variation (CV) and the mean for the tissue measures used in the analysis.

P-values for Breed* size interaction for RT-qPCR *SPP1* mRNA expression

A

Endometrium	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.0217	-	-	-
MS normal	0.1665	0.0035	-	-
MS smallest	0.9384	0.0246	0.2075	-

B

Placenta	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.1863	-	-	-
MS normal	0.1427	0.7953	-	-
MS smallest	0.1497	0.82	0.9757	-

C

Whole Uterus	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.7974	-	-	-
MS normal	0.7791	0.9691	-	-
MS smallest	0.6035	0.7785	0.8177	-

Table 6.9 Tables showing p-values from the pair wise comparison of least square means for the Breed * size interaction for *SPP1* mRNA expression level from RT-qPCR for LW and MS samples for endometrium (A), placenta (B) and whole uterus (C).

6.3.2. IHC

6.3.2.1. Location of SPP1 protein in whole utero-placental units by immunohistochemistry

SPP1 protein was localised by IHC in whole utero-placental units of LW-LR, LW, and MS foetuses. The results of this analysis are shown in Figure 6.4 for LW-LR crossbred and Figure 6.5 for LW and MS. In these figures a representative picture of each of the structures present in the section is presented, together with a control for each of them. Visual differentiation between breeds or between foetal sizes was not

possible. However, the main objective of this experiment was the spatial location in the endometrium, including GE and LE, and in the placenta. Visual examination of the pictures revealed that SPP1 protein was localised in endometrial GE towards the apical region of the cells compared to the expression towards the basal region in the LE cells. The visual examination of the placenta tissue was inconclusive, even though spots of expression were observed in some of the pictures. The difficulty of this differentiation increased with the auto fluorescence of the tissue, which was mainly blue, but green in part.

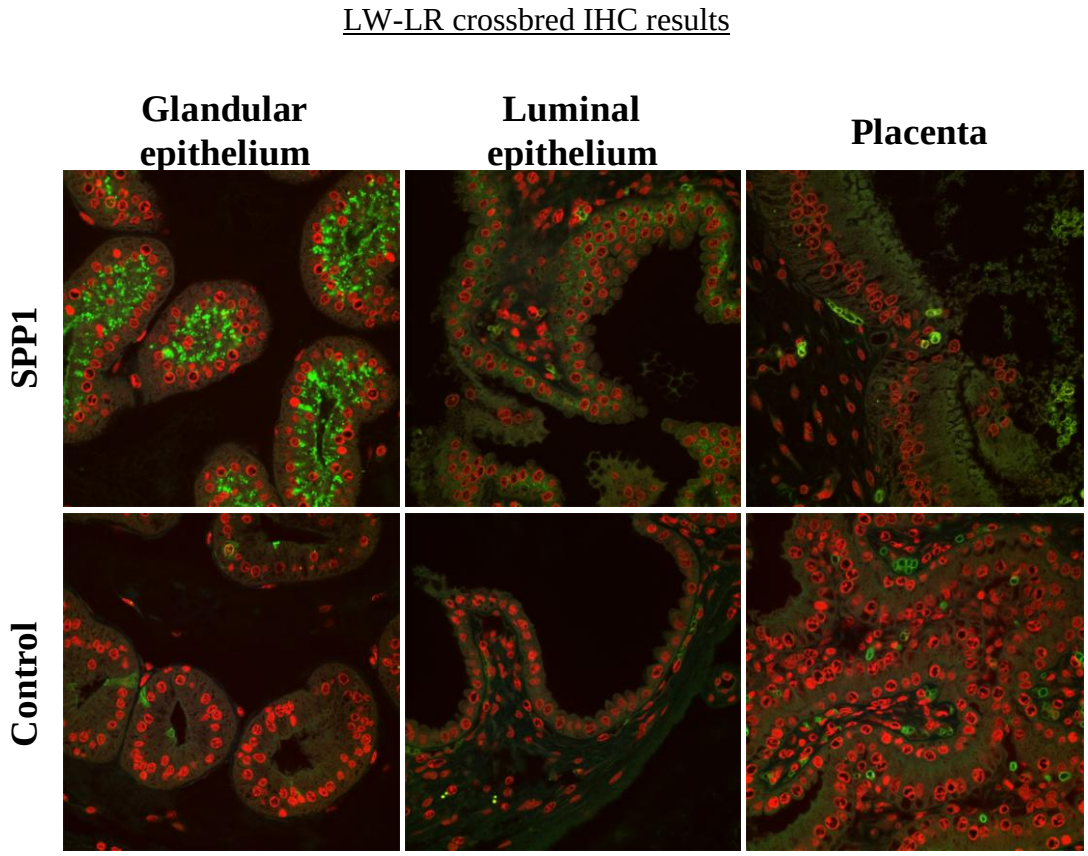

Figure 6.4 Detection of SPP1 protein in porcine whole utero-placental units from LW-LR crossbred by IHC staining. SPP1 protein was detected using a cocktail of IgG LF-124 and LF-166 as a primary antibody and Alexa Fluor as a secondary antibody, identifying SPP1 as green fluorescence under a confocal microscope at 60x magnification. The counterstaining with PI revealed the nucleus in red.

LW and MS IHC results

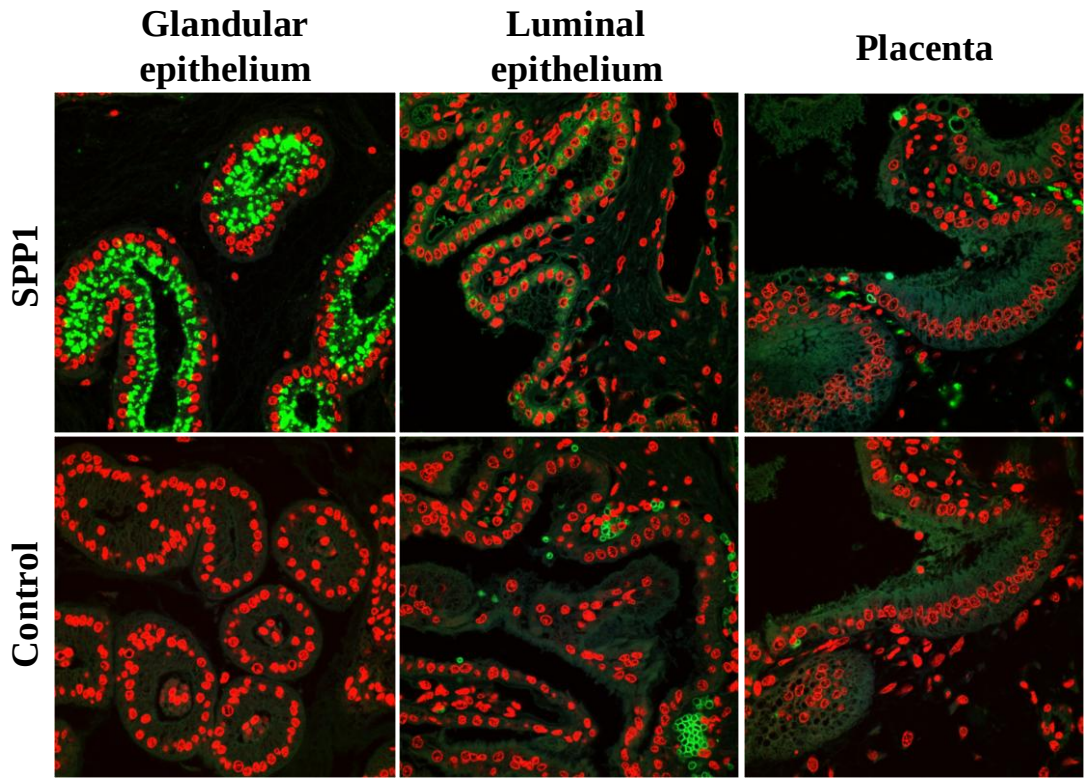

Figure 6.5 Detection of SPP1 protein in porcine whole utero-placental units from LW and MS purebred by IHC staining. SPP1 protein was detected using a cocktail of IgG LF-124 and LF-166 as a primary antibody and Alexa Fluor as a secondary antibody, identifying SPP1 as green fluorescence under a confocal microscope at 60x magnification. The counterstaining with PI revealed the nucleus in red.

6.3.2.2. Quantification of SPP1 protein in Large White-Landrace crossbred animals

The SPP1 protein was quantified in the three distinctive structures separately, as described in 6.2.10, and the results were analysed for differences between foetal sizes (6.2.11). The measure used was pixels of SPP1 per cell, calculated by dividing the total number of pixels of SPP1 protein in the area of interest by the number of nuclei in the same area. The mean values and standard errors for the three tissue structures for both foetal sizes were calculated and examination of these values indicates that

the expression of SPP1 protein in the normal foetus is higher than in the smallest ones for all the tissues.

The distribution of the results was checked for each tissue and the necessary transformations were performed, square root on GE, $\log_{10}$ (measure + 1) on LE, and $\log_{10}$ on placenta. In order to test the significance of the numerical differences, a general linear model of the protein expression results was performed, fitting size alone and size blocked by pig id to account for the common maternal environment. The results of this analysis, presented in Table 6.10, indicated that significant differences were found for all the tissues in the expression of SPP1 protein in smallest and normal foetuses when blocking for the maternal effect (P<.0001). When analysing the size alone, the only tissue where a significant effect was detected in the protein expression was the LE. However, the analysis where the gilt/sow was used to account for the maternal effect was considered to represent better the physiological implications and the objective of the study. The results, in which the maternal effect is included, indicate that SPP1 protein amount found in the smallest foetuses was lower than the one in the normal sized foetus, independently of the size and age of the gilt/sow and the stage of pregnancy.

	IHC Roslin		
	GE	**LE**	**Plac**
Model	<.0001	<.0001	<.0001
Size	0.5503	0.0198	0.427
Size (Pig id)	<.0001	<.0001	<.0001
R-Square	0.345	0.208	0.351
CV	49.118	63.484	92.647
Mean	11.021	5.250	2.847

Table 6.10 Results of the analysis of variance for the IHC analysis for quantification of SPP1 protein expression in the samples from LW-LR crossbred animals. The Table shows the p-values from an F-test for the model, size and pig id, and the R-square, the coefficient of variation (CV) and the mean for the tissue measures used in the analysis. GE, Glandular epithelium, LE, luminal epithelium and Plac, placenta.

6.3.2.3. Quantification of SPP1 protein in LW and MS purebred gilts

These samples, from LW and MS gilts, were analysed in the same way as the previous ones, being pixels of SPP1 per cell the final measure used for the analysis. In this case, the differences analysed were between breeds and between sizes. When analysing the breed differences, a general higher protein expression was observed in MS samples, regardless of the size of the foetus, compared with LW samples. It was observed, for both breeds, that the smallest foetuses had a higher protein expression than the normal ones, in all the tissues, except in placenta from MS foetuses, where the expression was higher in the normal foetuses. MS SPP1 expression tended to be higher than LW for both sizes.

The distribution of the final measure was verified and the values were transformed, as for the other data set, square root on GE, $\log_{10}$ (measure + 1) on LE and $\log_{10}$ on placenta. In order to test the significance of the observed differences, the different factors were analysed in a general linear model, blocking for gilt, and analysing the interaction between size and breed. The results, illustrated in Table 6.11, show a significant difference between breeds ($P<0.0001$). The interaction between breed and size including the blocking factor, gilt, shows a significant difference for all the tissues (Table 6.11). This interaction was analysed in more detail with the objective of checking all the combinations of size and breed (Table 6.12). The interactions verified for significant differences were: LW normal foetus with LW smallest foetus, MS normal foetus with MS smallest foetus, LW normal foetus with MS normal foetuses, and LW smallest foetus with MS smallest foetus.

The results for GE in the model (Table 6.11) indicated a significant difference in the interaction of breed and size when gilt was included as a blocking factor ($P=0.005$), but not when this blocking factor was excluded ($P=0.66$). As mentioned previously, this blocking factor is considered important due to the maternal effect affecting both foetuses in the same uterus, and thus, the model including this factor is considered more realistic/accurate. When the individual interactions (Table 6.12-A) were inspected, the expression in GE in the LW smallest foetuses was significantly higher

than in the normal sized foetuses (P=0.02). Also SPP1 protein level in MS normal foetuses was higher than LW foetuses of the same size (P=0.02).

In the LE tissue, the model indicated significant differences (P<.0001) for both size and breed, and its interaction including the blocking factor, gilt. When the individual interactions were analysed (Table 6.12**Error! Reference source not found.**-B), a significantly higher expression of SPP1 in the MS foetuses of both size was found comparing with the LW of each size, respectively (P=0.005 normal, P<.0001 smallest). Moreover, expression level in MS smallest foetus was significantly higher than in the normal foetuses in this breed. In the placenta tissue (Table 6.12-C), a significantly higher level of SPP1 protein was found in MS foetuses of both sizes when compared with the LW foetuses, respectively (P<.0001 normal, P=0.005 smallest). Also the LW foetuses were significantly different between them with higher expression in the smallest ones.

	IHC France		
	GE	**LE**	**Plac**
Model	<.0001	<.0001	<.0001
Breed	0.0064	<.0001	<.0001
Breed (Pig id)	<.0001	<.0001	<.0001
Size	0.0093	<.0001	0.601
Breed*Size	0.661	<.0001	0.024
Breed*Size(Pig id)	0.0051	<.0001	<.0001
R-Square	0.224	0.519	0.287
CV	36.080	33.986	71.712
Mean	16.040	14.127	3.193

Table 6.11 Results of the analysis of variance for the IHC results for SPP1 protein expression in the samples from LW and MS animals. The Table shows the p-values from an F-test for the model, breed blocked for pig id, size, and the interaction, and the R-square, the coefficient of variation (CV) and the mean for the tissue measures used in the analysis, SPP1 pixels per cell.

P-values for breed* size interaction for IHC SPP1 protein expression

A

Glandular epithelium	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.0237	-	-	-
MS normal	0.0253	0.92	-	-
MS smallest	0.0002	0.1039	0.1472	-

B

Luminal epithelium	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.4056	-	-	-
MS normal	0.0051	0.0444	-	-
MS smallest	<.0001	<.0001	<.0001	-

C

Placenta	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.0442	-	-	-
MS normal	<.0001	<.0001	-	-
MS smallest	<.0001	0.0054	0.2278	-

Table 6.12 Tables showing p-values from the pair wise comparison of least square means for the Breed * size interaction for SPP1 protein expression level from IHC for LW and MS samples for glandular epithelium (A), luminal epithelium (B) and placenta (C).

6.3.3. ISH

6.3.3.1. Probe development

An antisense and a sense probe for *SPP1* mRNA were prepared in the laboratory for the performance of ISH analysis in the tissues collected. For the preparation of the probe, RNA isolated from three samples from the two first pigs was selected as the mRNA template. As described previously, this RNA was reverse transcribed (6.2.4), and the resulting cDNA was used for the amplification of *SPP1* mRNA by PCR (5.2.13). The results of the three amplifications were inspected loading the samples in a gel (6.2.8), and running them together with a ladder. The results are presented in

Figure 6.6, where a fragment of the size of the amplification correspondent to the *SPP1* primers used was observed for the three samples.

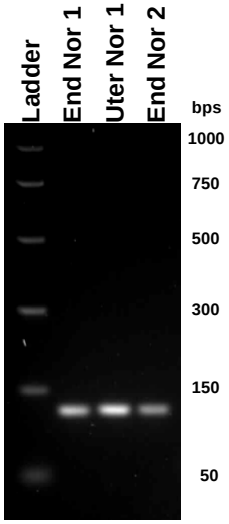

Figure 6.6 Gel for *SPP1* amplification for the preparation of the probe. Three samples were used, End Nor 1, endometrium normal pig1, Uter Nor 1, whole uterus normal pig1 and End, Nor 2, endometrium normal pig2. The samples were run together with a ladder.

The bands were excised from the gel, the cDNA was extracted as described in 6.2.9.1 and its concentration measured (Table 6.13). These fragments were used for a ligation into pGEMT easy vector (6.2.9.1). As described in Table 6.5, three different concentrations of insert were used in three different ligations with the corresponding controls. After the ligations were incubated overnight at 4ºC, JM109 high efficiency competent cells were transformed with a heat-shock to allow the vector to enter the cell and after a 1.5 hours incubation at 37ºC, the cells were plated in LB plates containing X-Gal, ampicillin, and IPGT, in duplicate for the samples and one for the controls. After the plates were incubated overnight at 37ºC, a white colony was picked up from each plate, resulting in two colonies for each of the samples and one colony for each of the controls. These colonies were grown in liquid LB overnight at 37ºC, and the DNA from the growth cells was purified with a miniprep system as described in the protocol (6.2.9.1).

Sample id	ng/µl	OD 260/280	OD 260/230
Endometrium normal Pig 1	29.26	1.62	0.05
Whole uterus normal Pig 1	6.08	3.67	0.01
Endometrium normal Pig 2	5.03	2.14	0.01

Table 6.13 Nanodrop results from SPP1 mRNA fragment purified from the gel. The values correspond to the concentrations of the inserts and the OD ratios.

Sample id	ng/µl	OD 260/280	OD 260/230
Endometrium normal Pig 1 Colony 1	98.81	1.86	1.99
Endometrium normal Pig 1 Colony 2	63.06	1.86	1.73
Whole uterus normal Pig 1 Colony 1	62.05	1.81	1.7
Whole uterus normal Pig 1 Colony 2	92.09	1.79	1.92
Endometrium normal Pig 2 Colony 1	61.55	1.84	1.75
Endometrium normal Pig 2 Colony 2	80.33	1.81	1.79
Transformation control	40	1.77	1.41
Positive control	57.21	1.8	1.66
Background control	53.62	1.81	1.73

Table 6.14 Concentrations of the samples after the miniprep purification.

An aliquot of the resulting vectors was digested with NotI enzyme which has restriction sites at both sites of the insert, and run together with an aliquot of its corresponding undigested vector in a gel, and the size of the resulting fragment was verified. The results, presented in Figure 6.7, indicated the presence of a unique vector with a small fragment (Uter Nor Set1 CutA), which size was verified in another gel with a ladder adequate to the size of the insert (Figure 6.8). Once the size was confirmed, the sample was sequenced and the resulting sequence was compared with the published sequence (NM-214023). Some of the cells preserved in Glycerol were plated in two LB plates containing ampicillin. After the overnight incubation at 37ºC, one defined colony was grown in 25 ml of liquid LB. Once the LB with the colony was incubated overnight, the vector was lysed from the cells and purified. The concentration of the vector was verified with the Nanodrop, showing a 76.5 ng/µl concentration. In order to prepare the antisense and sense, the sequence (vector + insert) for two different enzymes with cutting sites at only one side of the insert (Figure 6.9).

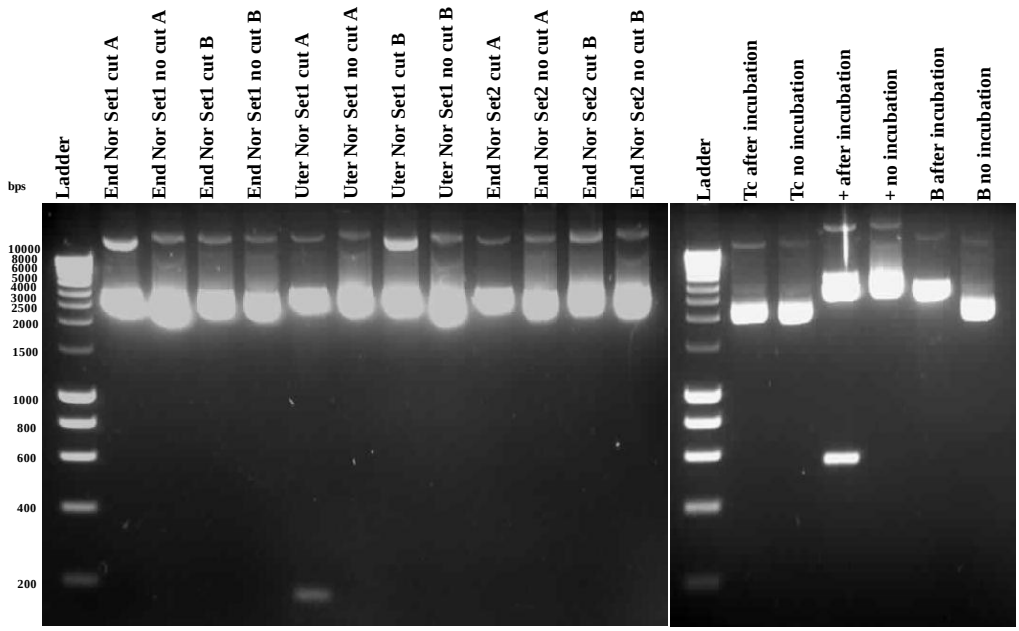

Figure 6.7 Gel of the vectors digested with NotI. Aliquots of vectors before and after digestion were run in the gel for all the samples. Also the controls were run in the gel. End, endometrium, Uter, uterus, Nor; normal, Tc, transformation control. Ladder of 10,000 bps.

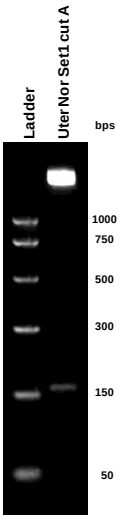

Figure 6.8 Gel corresponding to the vector with the correct fragment (Uter Nor Set1). The size of the fragment is indicated by the 1,000 bp ladder.

Figure 6.9 Sequence for both strands of the insert and part of the vector with RNA polymerase T7 and SP6. All the available enzyme restriction sites are indicated for both sequences. *SPP1* primers are indicted, as well as the polymerase transcription initiation sites.

Cutting with one enzyme in only one place will leave a linearised vector. Cutting with two different enzymes will give two different linearised vectors and in each one the insert will be amplified with one of the two different polymerases with transcription initiation sites in this vector. One of the requirements of the riboprobe kit was the absence of 3' overhangs, meaning that enzymes that form these cuts could not be used for the preparation of the templates. This eliminates PstI, SacI, and BstXI enzymes as candidates for the templates preparation. Other enzymes that were not allowed were the ones with two restriction sites in the vector, because they will cut the insert without polymerase transcription site. These enzymes were EcoRI, NotI, and BstZI.

In this way, each of the RNA polymerase produced a run-off transcript. Taking this into account, two different enzymes were chosen: SalI and NcoI. When cutting with SalI, the T7 polymerase site was next to the insert, and as the polymerase reads in direction 3' – 5' and synthesises in direction 5' - 3', the T7 polymerase synthesised the

sense template that is equal to the mRNA sequence. Thus, when the vector was cut with NcoI enzyme, the insert was next to the SP6 polymerase transcription initiation site and this polymerase will synthesise the antisense template, which was complementary to the mRNA and hybridised with it.

An aliquot of the purified plasmid was digested with each of the enzymes and the size and efficiency of the digestion was verified with a gel. As the vector looked partially digested, all the digest was run in the gel, and the bands were purified to ensure that only digested vector was used for the preparation of the riboprobe. The DNA was extracted from this purified vector, as described in 6.2.9.1. The concentration of the resulting DNA, template for the antisense and for the sense probe, was verified to assure that there was enough DNA. Then, these two templates were used to prepare the riboprobes with ^{35}S-UTP as indicated in 6.2.9.2. Each probe, together with the control, was used for the ISH in the whole utero-placental unit tissue slides as described in 6.2.9.3.

6.3.3.2. In situ *Hybridisation results*

Whole utero-placental unit sections were hybridised with the prepared probes, and after the corresponding incubation times and washes, the slides were developed and counter stained in Haematoxylin. Once the slides were dry, they were inspected under a bright and dark field microscope, in order to identify the tissue and observe the presence or absence of silver grain points.

Location of SPP1 *mRNA in Whole utero-placental units by ISH*

SPP1 mRNA was localised by *in situ* hybridisation in whole utero-placental units of LW-LR crossbred, LW and MS foetuses as described in 6.2.7. The results of this analysis are shown in Figure 6.10 for LW-LR crossbred and Figure 6.11 for LW and MS samples. In these Figures, as for the IHC results, a representative dark field antisense picture of each of the structures present in the section is shown together with its corresponding bright field picture.

Figure 6.10 Detection of *SPP1* mRNA in porcine whole uterus of pregnant gilts/sows at day 41 to 46 from a LW-LR crossbred by ISH analysis. For each tissue, a representative section with the correspondent bright-field and dark-field images with the antisense radiolabelled probe are shown. A representative section hybridised with the sense probe is also shown as a negative control of the hybridisation. 20x magnification. Sections were counterstained with haematoxylin in order to visualise the nucleus

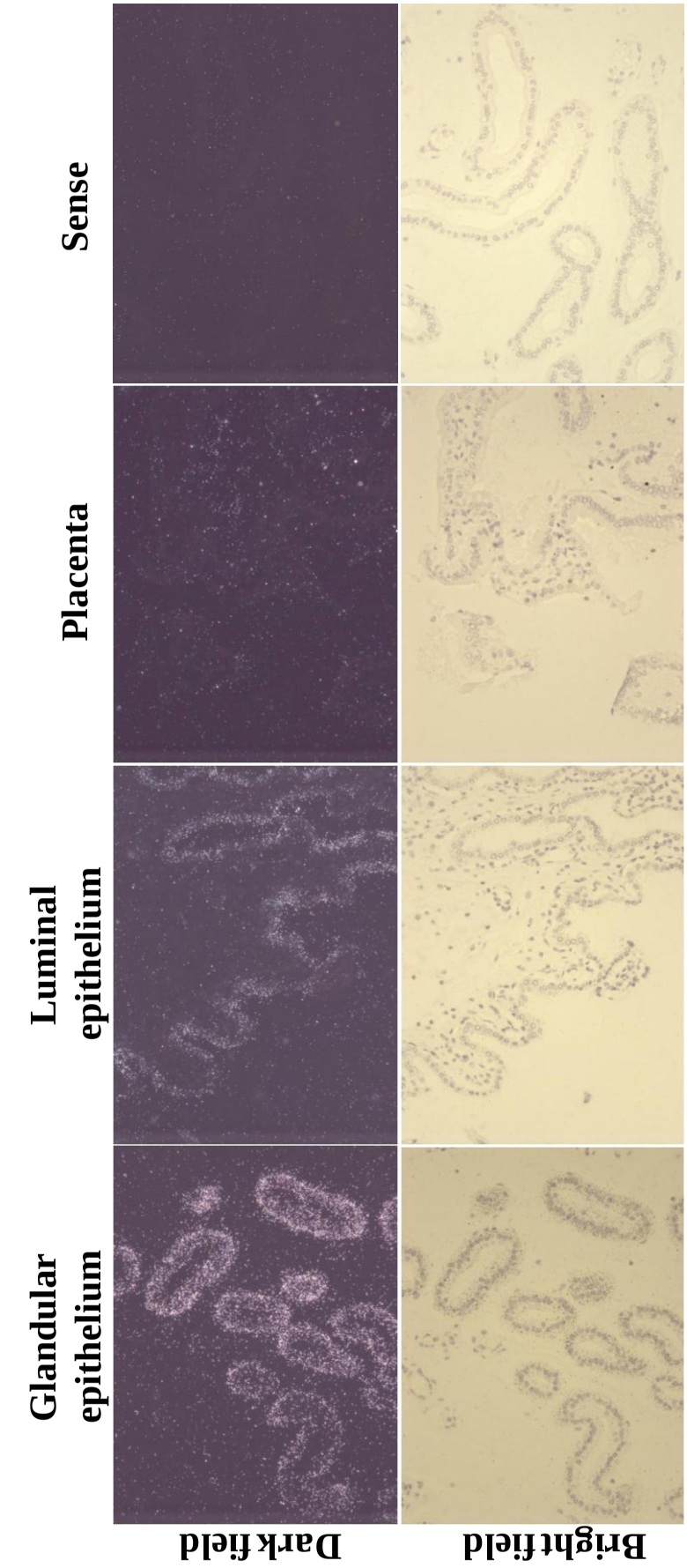

Figure 6.11 Detection of *SPP1* mRNA in porcine whole uterus of pregnant gilts at day 41-42 from LW and MS purebred animals. For each tissue, a representative section with the correspondent bright-field and dark-field images with the antisense radiolabelled probe are shown. A representative section hybridised with the sense probe is also shown as a negative control of the hybridisation. 20x magnification. Sections were counterstained with Haematoxylin in order to visualise the nucleus.

A representative dark and bright field sense picture for each sample set was also shown. However, for both samples, there is a clear *SPP1* mRNA expression in the glandular and luminal epithelium of the endometrium (GE and LE), with an overall distribution in the cell. As for the IHC, the results in the placenta are not very clear in most of the pictures, but some expression was evident. These pictures determined the spatial location of *SPP1* mRNA in the endometrium and placenta of pregnant pigs of LW and MS.

Quantification of SPP1 *mRNA in Large White-Landrace crossbred animals*

The *SPP1* mRNA was quantified in the three distinctive structures separately, as described in 6.2.10, and the results were analysed for differences between foetal sizes (6.2.11). The measure used was number of silver grains per cell, calculated dividing the total number of silver grains in the area of interest by the number of nuclei in the same area. The silver grains represented the *SPP1* mRNA. The mean values and standard errors for the three tissue structures for both foetal sizes indicated that the expression of *SPP1* mRNA in the normal foetus was higher than in the smallest ones for all the tissues.

The distribution of the sample was verified, and the measures were transformed as necessary, GE by square root, LE by $\log_{10}$, and placenta by square root. In order to test the significance of these differences, a general linear model of the mRNA expression results was performed, fitting size alone and size blocked by pig to account for the common maternal environment. The results of this analysis, presented in Table 6.15, indicated that significant differences were found for all the tissues in the expression of *SPP1* mRNA in smallest and the normal foetuses when blocking for the gilt ($P<.0001$). As shown by the graph, the mRNA level in the tissues supporting the normal sized foetuses was significantly higher than in those supplying the smallest foetuses ($P<.0001$).

	ISH Roslin		
	GE	**LE**	**Plac**
Model	<.0001	<.0001	<.0001
Size	0.0708	0.008	0.0001
Size(Pig id)	<.0001	<.0001	<.0001
R-Square	0.494	0.508	0.463
CV	28.403	60.208	24.220
Mean	2.269	0.431	1.507

Table 6.15 Results of the analysis of variance for the ISH for *SPP1* mRNA expression in the samples from LW-LR crossbred animals. The Table shows the p-values from an F-test for the model, size and size blocked by pig id, and the R-square, the coefficient of variation (CV) and the mean for the tissue measures used in the analysis, the normalised *SPP1* values.

Quantification of SPP1 *mRNA in LW and MS purebred gilts*

These samples, from LW and MS gilts, were analysed in the same way as the previous ones, being number of silver grains per cell the final measure used for the analysis. In this case, the differences were analysed between breeds and between sizes. The mean values and standard error for the final measures were calculated for both breeds and separating each size in both breeds. When analysing the breed differences, a general higher mRNA expression level was observed in LW samples, compared with MS samples. It was observed, for both breeds, that the differences between the sizes, respects to a higher or lower level of expression, vary between tissues and breeds.

The distribution of the raw data was verified, and the data points were transformed: GE by $\log_{10}$, LE by $\log_{10}$(measure + 1), and placenta by $\log_{10}$. In order to test the significance of these observed differences, transformed mRNA levels were studied in a general linear model, including breed and size as factors, blocking for gilt and analysing the interaction. The results, illustrated in Table 6.16, showed a significant difference between breeds ($P<0.0001$ for endometrium and $P=0.0002$ for placenta). The interaction between breed and size including the blocking factor, gilt, illustrated a significant difference for all the tissues (GE $P=0.007$, LE $P=0.002$, and placenta

P<.0001). The interactions (Table 6.17), as before, verified for significant differences were: LW normal foetus with LW smallest foetus, MS normal foetus with MS smallest foetus, LW normal foetus with MS normal foetuses and LW smallest foetuses with MS smallest foetuses.

For GE (Table 6.17-A), when the individual interactions were inspected, the expression in the LW smallest and normal sized foetuses was significantly higher than in the MS smallest and normal sized foetuses, respectively (P<.0001). For LE (Table 6.17-B), significant differences were found in all the interactions, between breeds and between sizes. Expression of *SPP1* mRNA was significantly higher in LW tissues of both sizes compared to the respective sizes in MS (P<.0001). There was also a significant difference in the size in each breed in this study but in different directions. The expression of LW smallest foetuses was significantly higher than their normal littermates (P=0.05). However, MS smallest foetuses presented a significantly lower level of mRNA than the MS normal foetuses (P=0.02).

	ISH France		
	GE	**LE**	**Plac**
Model	<.0001	<.0001	<.0001
Breed	<.0001	<.0001	<.0001
Breed(Pig id)	<.0001	<.0001	0.0002
Size	0.1855	0.0041	0.1344
Breed*Size	0.3323	0.7175	0.1145
Breed*Size(Pig id)	0.0072	0.0022	<.0001
R-Square	0.484	0.773	0.444
CV	26.391	16.474	43.878
Mean	0.836	0.775	0.437

Table 6.16 Results of the general linear model for the ISH results for *SPP1* mRNA expression in the samples from LW and MS animals. The Table shows the p-values from an F-test for the model, breed blocked for pig id, size, and the interaction, and the R-square, the coefficient of variation (CV) and the mean for the tissue measures used in the analysis, number of silver grains per cell.

Similarly, expression in placentas supplying the LW normal and smallest foetuses was significantly higher than the corresponding placentas from MS foetuses (P<.0001). In contrast to the LE, the mRNA in the MS normal foetuses was significantly lower than in their smallest litter mate (P=0.03).

P-values for Breed*size interaction for ISH *SPP1* mRNA expression

A

Glandular epithelium	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.7795	-	-	-
MS normal	<.0001	<.0001	-	-
MS smallest	<.0001	<.0001	0.1216	-

B

Luminal epithelium	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.0583	-	-	-
MS normal	<.0001	<.0001	-	-
MS smallest	<.0001	<.0001	0.0297	-

C

Placenta	LW normal	LW smallest	MS normal	MS smallest
LW normal	-	-	-	-
LW smallest	0.9684	-	-	-
MS normal	<.0001	<.0001	-	-
MS smallest	<.0001	<.0001	0.0378	-

Table 6.17 Tables showing p-values from the pair wise comparison of least square means for the Breed * size interaction for *SPP1* mRNA expression level from ISH for LW and MS samples for glandular epithelium (A), luminal epithelium (B) and placenta (C).

6.4. Discussion

This study represents the first time that SPP1 mRNA and protein levels were studied in relation with the effects of within-litter variation in foetal weight and breed differences in SPP1. Moreover, this analysis represents the continuation of a QTL

analysis described in this study in Chapter 3, allowing the confirmation of *SPP1* as a physiological candidate gene with distinctive expression in different breeds with marked prolificacy differences. The results presented here demonstrate the potential of SPP1 as a candidate gene for selection, as well as for a possible improvement in successful pregnancy and within-litter weight variation in pigs.

For this purpose, two different sets of samples were collected and mRNA and protein analyses performed. With the first set, from LW-LR crossbred animals, the main objective was to compare SPP1 mRNA and protein expression in reproductive tissues within-litter, comparing the smallest foetus and a normal sized foetus in the same litter. In the second samples, two different pig breeds were used; LW and MS purebred, as for the QTL analysis. This allowed the comparison not only within-litter but also between breeds. These two breeds are characterised for the different LS and foetal growth rate, as described previously. LW is a commercial breed, whereas MS is a prolific breed with carcass traits unsuited to commercial meat production.

The quantification of mRNA expression by RT-qPCR is a very precise technique and it was used in this study for the quantification of *SPP1* mRNA in endometrium, placenta and whole uterus. *SPP1* mRNA was also quantified by *in situ* hybridisation. With the exception of placental tissue, the relative abundance of *SPP1* in the different tissue types differed importantly between the qPCR and ISH analyses.

RT-qPCR is a technique characterised for its simplicity, sensitivity and specificity. ISH allows the specific location of the mRNA in the tissue on the selection of the tissue of interest, this represents a big difference between both techniques. However the specificity of ISH even if high is expected to be lower than in the RT-qPCR. Moreover, qPCR is the technique of choice nowadays for expression studies, such as validation of microarray results.

The problem that is presented here is the different distribution of tissues in both analyses, with only one tissue where a direct comparison was possible. However, these differences could also be explained by the number of samples analysed. The positive part of the ISH analysis, dismissing the lower sensitivity and specificity

compared to RT-qPCR, is the possibility of selecting the specific component of tissue to quantify, not possible in the RT-qPCR. The importance of this factor rise when the collection method is analysed. These tissues were collected manually and specifically the placenta tissue was peeled off from the rest of the tissues, which are the uterine tissues and the placenta is in direct contact with the endometrium. At the stage of pregnancy when the tissues were collected, the placenta was a fully differentiated tissue and the separation of the tissues was obvious. However, the placenta tended to break some times. Moreover in some of the sections of endometrium parts of placenta were found in a very low amount. There was not observation of endometrium tissue in placenta samples.

In a study by Miles *et al.* (2009) where the same tissues as in this study were collected but earlier in pregnancy, they observed the same invasion of placenta tissue where only endometrium was expected. Thus, this possibility in this study is not discarded, and was proved due to the visual observation of this mixture of tissues. Miles *et al.* (2009) described this fact as a cause of the absence of differences in mRNA at early stages of pregnancy.

However, in the present study, when comparing mRNA and protein studies, no correlation was found between the mRNA levels obtained from ISH, and the protein levels, but there was a correlation between the mRNA level from RT-qPCR and protein levels. The mRNA analyses complemented the protein analysis, to provide quantification and localisation of SPP1 mRNA and protein in the same tissues. However, the lack of correlation between mRNA and protein results probably reflects regulation processes occurring between the transcription of a gene and the translation of the encoded protein. Therefore, RT-qPCR was considered as the preferred technique in this study to quantify mRNA.

The differences in expression of SPP1 found in this study could be explained by its role in implantation, as well as in the maintenance of pregnancy. In these processes, SPP1 has a key role in the interaction between foetuses and maternal tissues. MS is a very prolific breed, due to its different strategies to maintain a large number of foetuses in the uterus, such as the protein secretion, synchrony in foetal growth, and

the efficiency of the placenta. This difference in LS is also observed in the samples collected in this study, where the higher mean LS is reported in the MS breed. Moreover, MS has been shown to have less within-litter variation in foetal weight through pregnancy when compared to commercial breeds, such as LW (Finch *et al.*, 2002). As a consequence of this uniformity, the difference between the smallest MS foetus in a litter and the normal one is lower than in other breeds, as observed in the samples used in this study (Table 6.3). This difference is a measure of the magnitude of growth retardation present in the smallest foetuses. Therefore, the lower retardation in foetal growth presented in MS and translated in larger LS, could explain the difference in the SPP1 expression direction when comparing the different breeds, such as protein levels in placenta from MS foetuses.

Therefore, the results found in this study suggest that there is an increase in the expression of SPP1 protein in tissues supplying the smallest foetuses, in order to compensate for the size and improve the growth and survival chance through an increase in attachment between the placenta and the uterus in the non-invasive implantation presented in pigs.

It is clear that further analysis are needed in larger samples, but the significant differences found in this study, not only within-litter but also between breeds, are really promising, and confirmed the important role of SPP1 in maintenance of pregnancy. Moreover, this study demonstrated, for the first time, that SPP1 can be used to measure the effectiveness of conceptus attachment associated with foetal growth and weight, not only to analyse growth within a litter but also between litters of diverse genetic background.

Chapter 7
General Conclusions

7.1. Motivation

Reproductive success is an important factor in livestock production, influenced by both genes and environment (Hume *et al.*, 2011). One of the challenges in this context is the improvement of reproductive traits, such as LS at a low cost, in order to improve the sustainability of the production system. However, although environmental factors have a high influence on reproduction, they are generally easy to control and do not constitute one of the biggest concerns (Prunier *et al.*, 2010). Nevertheless, nowadays molecular genetics approaches, such as marker assisted selection, are becoming more relevant for reproductive traits, as well as other economically important traits, as they allow improved performances, at a lower cost and for traits with low heritabilities (Williams, 2005; Spotter & Distl, 2006). Nowadays, breeding companies scan the animals for beneficial alleles to select for improvement in certain traits (www.thepigsite.com*)*.

Pigs are a good model for genetic studies, due to their short generation interval, early puberty and large LS, compared to other species; characteristics which make them ideal for studies requiring a large number of individuals, possibly from different generations (Rohrer *et al.*, 1999). Moreover, the pig is an important animal from two different points of view, i.e., for meat production and as a biomedical model. In the last decade, the production of pork has increased due to an increase in demand of this meat, in a market that is more and more demanding regarding the quality of the product. The other aspect, i.e., the use of the pig as a biomedical model, is related to the similarities (both genetic and physiological) between the human and the pig, which make the study of certain characteristics in the pig important (Meyers *et al.*, 2005; Cooper *et al.*, 2008).

A large number of studies have examined the characteristics and reproductive performance of the different pig breeds around the world in order to improve productivity in the swine industry. These studies have analysed the different factors affecting reproduction success in the different populations, showing that most of the reproductive traits, such as LS, OR and ES, have low heritability (Bennett &

Leymaster, 1989; Rothschild, 1996; Johnson *et al.*, 1999d). Moreover, another difficulty from an animal breeding perspective is that these traits can be recorded only on sexually mature females (Avalos & Smith, 1987). Thus, rapid improvement in reproduction through direct selection is difficult. However, many attempts have been made to improve such traits, with different degrees of success (Cunningham *et al.*, 1979; Bennett & Leymaster, 1989; Bennett & Leymaster, 1990; Lamberson *et al.*, 1991; Rathje *et al.*, 1997; Johnson *et al.*, 1999d; Cassady *et al.*, 2001; Mesa *et al.*, 2003).

The models that best describe the genetic control of reproduction traits, as with other so-called complex traits, assume that these traits are influenced by several genes each with an infinitesimal effect on the end trait. Therefore, it is unlikely that any one gene will be the key to selective breeding to improve these traits. One approach to address this problem is to use high density SNP genotypes to apply selection for the trait(s) of interest across the genome in so-called genomic selection. However, where the aim is to understand the molecular details of the genetic control of the trait of interest it is necessary to map, identify and characterise the genes responsible (Buske *et al.*, 2006a).

One of the traits showing more variation between populations is LS, which has low heritability and is a composite trait. The latter, therefore, suggests that there are different ways of increasing LS, such as by increasing OR, ES, UC and/or placental efficiency. It is indeed shown that direct selection for LS has been less successful than selection considering a combination of related traits (Bennett & Leymaster, 1990; Johnson *et al.*, 1999d; Ruiz-Flores & Johnson, 2001). All these traits have an influence on LS and present variation among breeds. This variation, both at the genetic and physiological level, indicates the possibility of improvement through selection. One of the breeds used in such studies is the Chinese MS breed, characterised by high prolificacy. It has been shown that the high prolificacy is based on increased PS, unrelated to the OR (Haley & Lee, 1993). Different factors have been related with this increase in PS, such as the smaller size of the piglets.

Moreover, the increased UC of this breed can be attributed to differences between MS and other breeds. The differences can be observed from the start of pregnancy. During early pregnancy, the complexity of the process is high. The embryos need to travel in the uterus and then elongate, and during this process oestrogen is secreted from the conceptus for maternal recognition of pregnancy. Western breeds express oestrogen in an embryo growth dependent manner, causing differences in the development of the embryos. However, MS follow a stage-dependent expression pattern for oestrogen, as well as a reduction in its expression (Pickard *et al.*, 2003). Moreover, the concentration of specific components of uterine secretions is reduced. These two factors help to reduce the competition between embryos thanks to the synchrony in growth, and, thus, the number of embryos lost. This strategy also affects the growth rate in MS, in which the piglets are more uniform in size, not only at birth but throughout pregnancy (Ashworth *et al.*, 1990a; Anderson *et al.*, 1993; Ford & Youngs, 1993; Youngs *et al.*, 1994; Wilson *et al.*, 1998; Wilson *et al.*, 1999). The distribution of the embryos in the uterus is uniform and they have similar size compared to Western breeds, such as LW and LR (Finch *et al.*, 2002). All these factors and the synchrony in growth between the developing embryos, reduces embryo mortality relative to the rates seen in other breeds. Another difference with the other breeds has been reported in later stages of pregnancy, when the foetuses need to grow rapidly. Most breeds show an increase in placental size to support this fast growth, whereas MS increase the placenta vascularity, without an increase in the size of the placenta (Biensen *et al.*, 1999). Also, due to the non-invasive nature of the pig placenta, the attachment of the placenta to the uterus is very important for the high rate of exchange required at this stage. Therefore, both strategies, i.e., the increased vascularity and the increase in attachment surface have been defended, and there is no evidence to say that one of them is not valid (Ford, 1997; Wilson *et al.*, 1998; Mesa *et al.*, 2003). As a result, MS piglets are smaller than those from other breeds at birth, but less variable in size, and this similarity reduces any competition between littermates. Therefore, the lower weight does not compromise the survival of these piglets.

Although, MS pigs are not good for the market, due to their low growth rate and high carcass fat content they are a valuable resource for studying the genetics of reproduction. In order to identify the gene(s) implicated in the strategy to reduce prenatal mortality, several crossbred studies have been performed. These studies use crosses between breeds with differing reproductive characteristics, in order to identify genomic regions affecting these characteristics. Nowadays, these studies have become more and more accessible, due to the lower cost of the techniques used. Thus, a large number of QTL have been mapped for these traits in the last decades. The next step is the identification and characterisation of the positional candidate genes discovered in the mapping studies (Buske *et al.*, 2006a).

7.2. Objectives

Many published QTL mapping studies are limited to a description of the mapped QTL and a comparison of the results from similar studies. Until recently, the identification of positional candidate genes for QTL in pigs has been limited by the low resolution of the mapping studies and by the absence of a fully annotated pig genome sequence. The number of QTL mapping studies and genome scans in which positional candidate genes have been characterised is low, as is the number of positional and physiological candidate genes found up to now that could be used for selection (Chapter 1-1.8).

As draft pig genome sequences have been released, the number of annotated genes has increased and it is easier to locate positional candidate genes. However, the function of these genes is not always known. Thus, most of the studies do not investigate further those candidate genes. In other studies, physiological candidate genes are analysed directly and searched for sequence variations in different populations, including sometimes characterisation at the mRNA and protein level. Results from these different approaches are related sometimes, but not always. However, the present study integrated both quantitative and physiological approaches, identifying a QTL with statistical support and searching for genes under the peak with a physiological role in reproduction.

With the fine mapping of SSC8 and the consequent confirmation and refining of a QTL for PS in this region, a positional and physiological candidate gene previously noted was confirmed in the present study, *SPP1*. Subsequently, mRNA and protein characterisation was studied in reproductive tissues, looking for the relation of this gene with foetal growth and weight differences in a litter and the differences in this gene between different breeds.

The general objective of a functional or characterisation study performed after a QTL study is the confirmation of the physiological function of the gene. Moreover, despite the increase in the number of genes mapped in the pig genome, most by sequence homology with the human genome, the function of these genes is often unknown. For this reason, an extensive study is necessary in order to confirm the gene as a physiological candidate gene. The improvement in functional studies allows large-scale expression studies, discovering numerous genes differentially expressed during pregnancy or between different breeds (Tuggle *et al.*, 2007; Fernandez-Rodriguez *et al.*, 2011). Nevertheless, individual characterisation of the genes in the tissues of interest is still warranted. For that, the expression needs to be located in a relevant tissue, where the function of this gene has a direct effect on reproduction. Once the location is confirmed, the pattern of expression needs to be characterised. As mentioned previously, there is a high variation in pigs, not only between breeds but also within populations. Thus, if individuals differ in reproductive traits, a different level of expression would be expected, if not between individuals from the same population, between individuals with different reproduction output and success, as in different breeds.

For all these reasons, the execution of a quantitative study for these traits and the mapping of candidate genes that could be characterised in the tissues of interest is very valuable. In the present study, *SPP1* was characterised in placental and endometrial tissues, in order to investigate the variations previously found in the genetic approach. For that, expression level and location of SPP1 mRNA and protein were compared not only in foetus of different size occupying the same uterus but also in different breeds with distinctive reproduction success.

7.3. Relevant findings

As a result of the analyses described in Chapter 2 and Chapter 3, 13 QTL were mapped in The Roslin LWxMS experimental crosses, of which only two had been previously mapped in this population and three, on the same chromosome, had not been mapped previously in other studies. Some candidate genes were described as positional candidate genes in these regions. However, due to the size of the QTL regions, the number of positional candidate genes is large. In this study only candidate genes with a physiological function in reproduction were discussed.

The results from the fine mapping analysis of SSC8 show that by increasing the marker density across a previously mapped QTL can not only improve the resolution with which the QTL is mapped but also change the perception of the number of QTL in the region (Chapter 3). In this study, three QTL were mapped on that chromosome, one with effects on PS and two co-located with effects on LS and TBA, respectively. In the earlier analysis (King *et al.*, 2003), the LS and PS QTL were not resolved and appeared co-located. Despite these changes in the estimated positions of these QTL, *SPP1* remained a positional candidate gene as it was located under the peak for PS on SSC8. The separation of the PS and LS/TBA QTL at 105 cM and 124 cM, respectively, on SSC8 suggests that there may be more than one gene in this region with effects on the end trait – TBA. The region of the QTL for LS and TBA, was inspected for genes with functions in reproduction.

The plots for the PS, TBA and LS QTL have a striking twin-peak appearance. The PS QTL has two peaks, even if only one position is significant and the LS QTL also has two peaks, one of which was significantly lower than the other. Two-QTL models were tested for this chromosome, but there was not enough evidence to support the presence of a second QTL. When either the QTL at 105 cM or 124 cM was fitted as background effect, QTL effects at the remaining position could be detected but without strong statistical support. Thus, although the two-QTL model does not provide a better fit to the data than the one-QTL model, there is not enough evidence to discard a second QTL in this region for either trait, PS and LS. The

importance of this region of SSC8 for reproduction was confirmed and there are a number of credible positional candidate genes that merit further investigation.

Positional candidate genes were also identified for QTL located across the genome and with effects on a range of reproductive traits (Chapter 2). Similarly, positional candidate genes were identified for regions with effects on reproductive traits and found through SNP associations in the genome-wide association study of a commercial population (Chapter 4). The QTL mapping in the experimental crosses and the GWAS study in the commercial population provided opportunities to study the genetic control of losses during pregnancy. The laparoscopy recordings of ovulations and data for the subsequent litters allowed estimates of PS in the experimental crosses. On the other hand, stillbirths and mummified piglets were recorded in the commercial population. Unfortunately, there was no overlap between the two studies in the loci identified for the key economic trait which was recorded in both populations (TBA/LS).

SPP1 was selected as the candidate gene for further analysis. This gene is located within the QTL for PS and associations have been reported between *SPP1* polymorphisms and number of piglets weaned and number alive 21 days after birth (Korwin-Kossakowska *et al.*, 2002). SPP1 has a well known role in implantation and maintenance of pregnancy and its location in endometrium and placenta tissues, as well as its regulation, have been previously defined. Finally, at a practical level the reagents required for analysis of this gene and the SPP1 protein were available.

As described in Chapter 5 and Chapter 6, two distinct sets of samples were collected during this study, to investigate the variation of SPP1 expression between foetuses of different sizes and between breeds. The first sample, including 9 LW-LR crossbred pigs, presented variation between animals in stage of pregnancy and parity (first or second), with different weights among gilts/sows. The second sample comprised 5 LW and 4 MS purebred gilts collected in France at the same stage of pregnancy and parity. Although the number of individuals per class was small in the latter group, this sample was more homogeneous.

The analysis presented here is the first to compare the expression of SPP1 between tissues supplying foetuses of different sizes (i.e., the smallest and a normal-sized) occupying the same uterus, as well as comparison between breeds with different reproductive characteristics. The two experiments (1. LW-LR gilts/sows; 2. LW *vs* MS) were analysed separately, due to the differences between them. In the first group, although *SPP1* mRNA expression was higher in endometrium associated with the smallest foetuses and in the placenta of normal-sized foetuses, these differences in expression were not statistically significant. However, there were significantly higher levels of SPP1 protein in all the sampled tissues associated with normal-sized foetuses. In contrast in the other set of samples, a general higher expression of both mRNA and protein was found in the tissues associated with MS foetuses compared to LW ones. In the comparison of expression in relation to the size of the foetuses, in general mRNA was higher in tissues associated with normal-sized foetuses for both breed. In contrast, in general higher levels of protein were observed in tissues associated with the smallest foetuses. However, in the MS animals *SPP1* mRNA levels in endometrium associated with the normal-sized foetuses was lower than that from the smallest ones and SPP1 protein levels in placenta from the smallest foetuses was lower than that from normal-sized foetuses.

SPP1 expression in the different stages of pregnancy, as well as its regulation by oestrogen and progesterone have been previously studied (Garlow *et al.*, 2002; White *et al.*, 2005) (Chapter 1-1.8.4). Therefore, the main concern in this study was the difference between breeds and the function of SPP1 in attachment in relation with foetal growth. The SPP1 transcripts and protein in the endometrium is, in part, maternally regulated, however it is the oestrogen from the conceptus which activates its expression and consequently its secretion from the LE. However, one of the advantages of pigs is their non-invasive type of implantation, where the only contact between the maternal and the foetal side is the apposition of the placenta and the endometrium (Burghardt *et al.*, 2002), and the interface of both membranes is where all exchange and transport of nutrients take place (Garlow *et al.*, 2002). It is in this interface where SPP1 fulfils its function as an adhesion molecule. MS foetuses do not suffer from the within-litter birth weight variation seen in other breeds, which in

this study was hypothesised to be a result of the attachment efficiency and as a function of SPP1 (Finch *et al.*, 2002). The higher levels of mRNA and protein found in the MS foetuses may indicate that SPP1 has a function in this efficiency in MS pigs. In the LW pigs the levels of SPP1 protein associated with the smallest foetuses is higher than for normal-sized foetuses, although the mRNA levels are lower. These higher levels of SPP1 protein associated with the smallest LW foetuses could be seen as a compensatory expression for the lower weight, and an effort to maintain that foetus until term. In MS the differences are in the mRNA in the endometrium and in the protein in the placenta. In this case, the higher level of mRNA in the smallest foetuses in the endometrium can be seen as a compensatory effect coming from the maternal side, since the expression of protein in the placenta is higher in the normal foetuses.

7.4. Conclusions

At this stage, the novelty of this study should be borne in mind. The complete QTL analysis was performed, including the goal, i.e., finding a candidate gene and analysing the expression in relevant tissues to verify the implications in reproduction. Thus, SSC8 QTL for PS was confirmed, together with another QTL position for LS and TBA. Ten more QTL were mapped across the genome with effects on reproduction, both confirming QTL found by others and finding new ones. Also some positional and physiological candidate genes were assigned to these QTL. The analysis of SSC8 was complemented with the characterisation of *SPP1*, a positional and physiological candidate gene. For the first time, SPP1 expression was analysed in the prolific MS breed, and compared with a commercial Western breed, as well as analysing SPP1 as a measure for foetal growth in foetuses of different sizes within-litter. The results of this analysis showed variation of SPP1 expression between breeds, in agreement with the QTL results, with higher expression in tissues supplying MS foetuses. Also differences in SPP1 were found between feto-placental units of different size occupying the same uterus. The results from the scan of The Roslin population, together with the SNP association analysis in a multi-line population, demonstrated the large number of genes that are positional candidate

genes with the increasing annotation of genes in the pig genome, whose function needs to be explored in relation to reproduction.

Using this gene for selection could generate a financial benefit in the pork industry. With the figures collected in 2010, when there were 424,500 heads in breeding herds and 22.1 piglets were reared per sow per year (BPEX annual report 2010-2011), the use of the QTL for PS mapped in this study could provide a £13m benefit due to an increase of around 0.7 piglets per sow per year. However, for a long term profit a reduction on the breeding sows stock would be recommended. Thus, a reduction of 12,500 sows would produce the same number of annual reared piglets and a reduction in part of the production cost.

7.5. Implications and future work

The current and projected future increased demand for pork meat in the market will require an increase in pig production (www.usda.gov). Therefore, an increase in production with an associated reduction in cost is of great interest for the porcine industry. Reproduction performance is one of the factors affecting the efficiency of the system and an increase in LS could increase the profitability of the sector. MAS using markers based on the genes determining differences in reproductive performance or in tight linkage disequilibrium with such causal genetic variants could contribute to achieving this goal (Meyers *et al.*, 2005; Spotter & Distl, 2006). Nowadays the number of characterised genes available for selection is low and use of linked genetic markers give diverse responses in different populations. For these reasons this area of research merits further investigation.

Some advance has been made by the study presented here, but due to the number of genes affecting these traits more needs to be done. QTL studies have mapped many regions affecting these traits, and some of them merit more analysis (Onteru *et al.*, 2009) (Chapter 1-1.8). Recent and anticipated improvements in annotation of the draft pig genome sequence (Archibald *et al.*, 2010) should be inspected in these

regions in order to identify positional candidate genes for subsequent physiological analyses.

New genomic tools and information could improve the effectiveness of the searches and the initial assessment of positional candidate genes. For example, the 60K SNP chip (Ramos *et al.*, 2009) offers the opportunity to analyse the whole genome in a large number of animals in a fast and reliable way, without the time consuming task of genotyping microsatellites individually. However there are also problems associated with this technique. The main concern is in its interpretation, due to the early stage of these genome-wide analyses. Moreover, due to the large number of SNPs analysed simultaneously, a larger number of animals is required than in QTL mapping in experimental crosses, in order to exploit the power of high density SNP genotyping for high resolution genetic analyses. Significance thresholds also need to be modified appropriately in order to minimise the number of false positives that can arise from multiple testing without being over stringent and risking false negatives. Nevertheless such genome-wide association studies in commercial populations have the potential to identify large number of candidate regions and trait genes.

The genome-wide analysis with the SNP chip performed in this study requires further analysis. The genes located in the associated regions need extensive analysis in order to study their function in reproduction. Beside, this analysis could benefit greatly from the addition of more animals. However, with the available information further analysis could be done, such as the performance of permutations in order to calculate a more precise p-value, and the separate inspection of each parity in a different analysis. The problem when analysing the results of these analysis is the lack of studies, to date, with which to compare the results.

The expression analyses could also benefit from including samples from more animals and developmental stages. However, there is a cost-time factor implicated in the increase in the number of animals. It would be important to confirm the functional analyses described here for SPP1 in other populations to confirm the effect and the use of this gene for MAS. Meanwhile, however, the tissues collected in this study can be used for the characterisation of other candidate genes.

The integration of microarray analyses with the QTL analyses is of great advantage for the designation of physiological candidate genes. This represents, in part, the other approach for the location of candidate genes starting from the function, finding physiological candidate genes that are differentially expressed in pigs of different reproductive performance, including diverse prolific and less prolific breeds. These analyses have already been used to verify the function of positional candidate genes from QTL analysis (Noguera *et al.*, 2009; Fernandez-Rodriguez *et al.*, 2011). The analysis presented here could also benefit from this verification, despite the low number of microarrays in reproductive tissues. This comparison is only part of the confirmation for a physiological candidate gene and further studies are required like the ones performed in the present study.

Gene expression transcript-level profiles being generated by colleagues at The Roslin Institute for multiple pig tissues and cell types using microarrays may also be useful in screening positional candidate genes. Positional candidate genes, which are also expressed in relevant tissues and at relevant developmental stages, could be prioritised for further functional analysis. Improvements in the annotation of the pig genome sequence would also allow the identification of non-protein coding positional candidate genes, including microRNA and other non-coding RNA genes.

The results from the characterisation analysis are very promising and demonstrated that the strategy for MS prolificacy has an important molecular background that can be exploited. The MS strategy is to increase the surface area of endometrial:placental attachment when the rapid growth of the foetus makes the uterine space limiting. In this way MS pigs increase UC which results in increased LS. Thus, higher levels of SPP1, as part of the attachment process in pigs, in MS may contribute to the higher efficiency in this breed. Another factor that plays an important role in the efficiency is placental vascularity, which is higher in MS when compared with other breeds and that contribute to the reduction in the increase of placenta size observed in other breeds during pregnancy. However, this factor seems to be determined in part by the conceptus. SPP1 expression is activated by conceptus oestrogen as a pregnancy indicator, but thereafter its expression is maintained by progesterone released from the corpora lutea. Thus, SPP1 is maternally controlled after activation and it is part of

the histotroph. However, the results presented here indicate a possible local regulation of SPP1 in the placenta by the conceptus. Therefore, SPP1 expression, location and regulation, in part determined by maternal effects, together with the location of the *SPP1* gene within a QTL which was mapped by analysis of maternal genotypes, make *SPP1* a good positional and physiological candidate gene and a possible marker for selection in pigs.

References

Alexander,L.J., Troyer,D.L., Rohrer,G.A., Smith,T.P., Schook,L.B., & Beattie,C.W. (1996). Physical assignments of 68 porcine cosmid and lambda clones containing polymorphic microsatellites. *Mamm Genome*, 7(5), 368-372.

Allan,M.F., Thallman,R.M., Cushman,R.A., Echternkamp,S.E., White,S.N., Kuehn,L.A., Casas,E., & Smith,T.P. (2007). Association of a single nucleotide polymorphism in SPP1 with growth traits and twinning in a cattle population selected for twinning rate. *J Anim Sci.*, 85(2), 341-347.

Anderson,L.H., Christenson,L.K., Christenson,R.K., & Ford,S.P. (1993). Investigations into the control of litter size in swine: II. Comparisons of morphological and functional embryonic diversity between Chinese and American breeds. *J Anim Sci.*, 71(6), 1566-1571.

Anderson,L.L. (1978). Growth, protein content and distribution of early pig embryos. *Anat Rec.*, 190(1), 143-153.

Anderson,L.L. (1966). Pituitary-ovarian-uterine relationships in pigs. *J Reprod Fertil Suppl.*, 1, 21-32.

Andersson,L. & Georges,M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. *Nat Rev Genet.*, 5(3), 202-212.

Apparao,K.B., Illera,M.J., Beyler,S.A., Olson,G.E., Osteen,K.G., Corjay,M.H., Boggess,K., & Lessey,B.A. (2003). Regulated expression of osteopontin in the peri-implantation rabbit uterus. *Biol Reprod.*, 68(5), 1484-1490.

Apparao,K.B., Murray,M.J., Fritz,M.A., Meyer,W.R., Chambers,A.F., Truong,P.R., & Lessey,B.A. (2001). Osteopontin and its receptor alphavbeta(3) integrin are coexpressed in the human endometrium during the menstrual cycle but regulated differentially. *J Clin Endocrinol Metab.*, 86(10), 4991-5000.

Arai,A., Nosaka,Y., Kanda,E., Yamamoto,K., Miyasaka,N., & Miura,O. (2001). Rap1 is activated by erythropoietin or interleukin-3 and is involved in regulation of beta1 integrin-mediated hematopoietic cell adhesion. *J Biol Chem.*, 276(13), 10453-10462.

Archibald,A.L., Bolund,L., Churcher,C., Fredholm,M., Groenen,M.A., Harlizius,B., Lee,K.T., Milan,D., Rogers,J., Rothschild,M.F., Uenishi,H., Wang,J., & Schook,L.B. (2010). Pig genome sequence--analysis and publication strategy. *BMC Genomics*, 11, 438.

Archibald,A.L., Haley,C.S., Brown,J.F., Couperwhite,S., McQueen,H.A., Nicholson,D., Coppieters,W., Van de,W.A., Stratil,A., Wintero,A.K., & . (1995). The PiGMaP consortium linkage map of the pig (Sus scrofa). *Mamm Genome*, 6(3), 157-175.

Ashworth,C.J. (2006). Reproduction. InI.Kyriazakis & C.T.Whittemore (Eds.), *Whittemore's Science and Practice of Pig Production* (pp. 104-147). Blackwell Publishing.

Ashworth,C.J., Finch,A.M., Page,K.R., Nwagwu,M.O., & McArdle,H.J. (2001). Causes and consequences of fetal growth retardation in pigs. *Reprod Suppl.*, 58, 233-246.

Ashworth,C.J., Fliss,M.F., & Bazer,F.W. (1990a). Evidence for steroid control of a putative angiogenic factor in the porcine uterus. *J Endocrinol.*, 125(1), 15-19.

Ashworth,C.J., Haley,C.S., Aitken,R.P., & Wilmut,I. (1990b). Embryo survival and conceptus growth after reciprocal embryo transfer between Chinese Meishan and Landrace x Large White gilts. *J Reprod Fertil.*, 90(2), 595-603.

Ashworth,C.J., Haley,C.S., & Wilmut,I. (1992). Effect of regumate on ovulation rate, embryo survival and conceptus growth in Meishan and Landrace × Large White gilts. *Theriogenology*, 37(2), 433-443.

Ashworth, C. J., McEvoy, T. G., Rooke, J. A., and Robinson, J. J. Nature, Nurture and the Quality of Life. 2-6. 2004. SAC Outlook 2004.
Ref Type: Report

Ashworth,C.J., Ross,A.W., & Barrett,P. (1998). The use of DNA fingerprinting to assess monozygotic twinning in Meishan and Landrace x large white pigs. *Reprod Fertil Dev.*, 10(6), 487-490.

Aulchenko,Y.S., de Koning,D.J., & Haley,C. (2007a). Genome-wide Rapid Association Using Mixed Model and Regression: A Fast and Simple Method For Genome-wide Pedigree-Based Quantitative Trait Loci Association Analysis. *Genetics*, 177(1), 577-585.

Aulchenko,Y.S., Ripke,S., Isaacs,A., & van Duijn,C.M. (2007b). GenABEL: an R library for genome-wide association analysis. *Bioinformatics*, 23(10), 1294-1296.

Avalos,E. & Smith,C. (1987). Genetic improvement of litter size in pigs. *Anim Prod.*, 44, 153-164.

Bailey,D.W., Dunlap,K.A., Frank,J.W., Erikson,D.W., White,B.G., Bazer,F.W., Burghardt,R.C., & Johnson,G.A. (2010). Effects of long-term progesterone on developmental and functional aspects of porcine uterine epithelia and vasculature: progesterone alone does not support development of uterine glands comparable to that of pregnancy. *Reproduction*, 140(4), 583-594.

Barr,A.J., Watson,S.P., Bernal,A.L., & Nimmo,A.J. (1991). The presence of NK3 tachykinin receptors on rat uterus. *Eur J Pharmacol.*, 203(2), 287-290.

Baxter,E.M., Jarvis,S., D'Eath,R.B., Ross,D.W., Robson,S.K., Farish,M., Nevison,I.M., Lawrence,A.B., & Edwards,S.A. (2008). Investigating the behavioural and physiological indicators of neonatal survival in pigs. *Theriogenology*, 69(6), 773-783.

Bazer,F.W. (1992). Mediators of maternal recognition of pregnancy in mammals. *Proc Soc Exp Biol Med.*, 199(4), 373-384.

Bazer,F.W. & Spencer,T.E. (2005). Reproductive biology in the era of genomics biology. *Theriogenology*, 64(3), 442-456.

Bazer,F.W., Spencer,T.E., Johnson,G.A., Burghardt,R.C., & Wu,G. (2009). Comparative aspects of implantation. *Reproduction*, 138(2), 195-209.

Bazer,F.W., Thatcher,W.W., Martinat-Botte,F., & Terqui,M. (1988). Conceptus development in large white and prolific Chinese Meishan pigs. *J Reprod Fertil.*, 84(1), 37-42.

Bee,G. (2007). Birth Weight of Litters as a source of Variation in postnatal growth, and Carcass and Meat Quality. *Advances in Pork Production*, 18, 191-196.

Beeckmann,P., Moser,G., Bartenschlager,H., Reiner,G., & Geldermann,H. (2003). Linkage and QTL mapping for Sus scrofa chromosome 8. *J Anim Breed Genet.*, 120, 66-73.

Bennett,G.L. & Leymaster,K.A. (1990). Genetic implications of a simulation model of litter size in swine based on ovulation rate, potential embryonic viability and uterine capacity: I. Genetic theory. *J Anim Sci.*, 68(4), 969-979.

Bennett,G.L. & Leymaster,K.A. (1989). Integration of ovulation rate, potential embryonic viability and uterine capacity into a model of litter size in swine. *J Anim Sci.*, 67(5), 1230-1241.

Bidanel,J.P., Caritez,J.C., & Legault,C. (1990). Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. II. Growth before weaning and growth of females during the growing and reproductive periods. *Genet Sel Evol.*, 22(4), 431-445.

Bidanel,J.P., Caritez,J.C., & Legault,C. (1989). Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. *Genet Sel Evol.*, 21(4), 507-526.

Bidanel,J.P., Milan,D., Iannuccelli,N., Amigues,Y., Boscher,M.Y., Bourgeois,F., Caritez,J.C., Gruand,J., Le,R.P., Lagant,H., Quintanilla,R., Renard,C., Gellin,J., Ollivier,L., & Chevalet,C. (2001). Detection of quantitative trait loci for growth and fatness in pigs. *Genet Sel Evol.*, 33(3), 289-309.

Bidanel,J.P., Rosendo,A., Iannuccelli,N., Riquet,J., Gilbert,H., Caritez,J.C., Billon,Y., Amigues,Y., Prunier,A., & Milan,D. (2008). Detection of quantitative trait loci for teat number and female reproductive traits in Meishan x Large White F2 pigs. *animal*, 2(06), 813-820.

Biensen,N.J., Wilson,M.E., & Ford,S.P. (1999). The impacts of uterine environment and fetal genotype on conceptus size and placental vascularity during late gestation in pigs. *J Anim Sci.*, 77(4), 954-959.

Biensen,N.J., Wilson,M.E., & Ford,S.P. (1998). The impact of either a Meishan or Yorkshire uterus on Meishan or Yorkshire fetal and placental development to days 70, 90, and 110 of gestation. *J Anim Sci.*, 76(8), 2169-2176.

Blomberg,L.A. & Zuelke,K.A. (2004). Serial analysis of gene expression (SAGE) during porcine embryo development. *Reprod Fertil Dev.*, 16(1-2), 87-92.

Bolet,G., Botte,F., Locatelli,A., Gruand,J., Terqui,M., & Berthelot,F. (1986). Components of prolificacy in hyperprolific Large White sows compared with the Meishan and Large White breeds. *Genet Sel Evol.*, 18(3), 333-342.

Bradford,G.E. (1968). Selection for litter size in mice in the presence and absence of gonadotropin treatment. *Genetics*, 58(2), 283-295.

Braunschweig,M.H., Paszek,A.A., Weller,J.I., Da,Y., Hawken,R.J., Wheeler,M.B., Schook,L.B., & Alexander,L.J. (2001). Generation and exploration of a dense genetic map in a region of a QTL affecting corpora lutea in a Meishan x Yorkshire cross. *Mamm Genome*, 12(9), 719-723.

Burghardt,R.C., Johnson,G.A., Jaeger,L.A., Ka,H., Garlow,J.E., Spencer,T.E., & Bazer,F.W. (2002). Integrins and extracellular matrix proteins at the maternal-fetal interface in domestic animals. *Cells Tissues Organs*, 172(3), 202-217.

Buske,B., Brunsch,C., Zeller,K., Reinecke,P., & Brockmann,G. (2005). Analysis of properdin (BF) genotypes associated with litter size in a commercial pig cross population. *J Anim Breed Genet.*, 122(4), 259-263.

Buske,B., Sternstein,I., & Brockmann,G. (2006a). QTL and candidate genes for fecundity in sows. *Anim Reprod Sci.*, 95(3-4), 167-183.

Buske,B., Sternstein,I., Reissmann,M., & Brockmann,G. (2006b). Detection of novel single-nucleotide polymorphisms (SNPs) in the CYP21 gene and association analysis of two SNPs for CYP21 and ESR2 with litter size in a commercial sow population. *J Anim Breed Genet*, 123(5), 343-348.

Bustin,S.A., Benes,V., Garson,J.A., Hellemans,J., Huggett,J., Kubista,M., Mueller,R., Nolan,T., Pfaffl,M.W., Shipley,G.L., Vandesompele,J., & Wittwer,C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem.*, 55(4), 611-622.

Bustin,S.A., Benes,V., Nolan,T., & Pfaffl,M.W. (2005). Quantitative real-time RT-PCR--a perspective. *J Mol Endocrinol.*, 34(3), 597-601.

Campbell,E.M., Nonneman,D.J., Kuehn,L.A., & Rohrer,G.A. (2008). Genetic variation in the mannosidase 2B2 gene and its association with ovulation rate in pigs. *Anim Genet.*, 39(5), 515-519.

Campbell,E.M., Nonneman,D.J., & Rohrer,G.A. (2003). Fine mapping a quantitative trait locus affecting ovulation rate in swine on chromosome 8. *J Anim Sci.*, 81(7), 1706-1714.

Campbell,S., Swann,H.R., Seif,M.W., Kimber,S.J., & Aplin,J.D. (1995). Integrins and adhesion molecules: Cell adhesion molecules on the oocyte and preimplantation human embryo. *Hum Reprod*, 10(6), 1571-1578.

Candenas,M.L., Magraner,J., Armesto,C.P., Anselmi,E., Nieto,P.M., Martin,J.D., Advenier,C., & Pinto,F.M. (2001). Changes in the expression of tachykinin receptors in the rat uterus during the course of pregnancy. *Biology of reproduction*, 65(2), 538-543.

Carson,D.D., Lagow,E., Thathiah,A., Al-Shami,R., Farach-Carson,M.C., Vernon,M., Yuan,L., Fritz,M.A., & Lessey,B. (2002). Changes in gene expression during the early to mid-luteal (receptive phase) transition in human endometrium detected by high-density microarray screening. *Mol Hum Reprod.*, 8(9), 871-879.

Carson,D.D., Bagchi,I., Dey,S.K., Enders,A.C., Fazleabas,A.T., Lessey,B.A., & Yoshinaga,K. (2000). Embryo Implantation. *Dev Biol.*, 223(2), 217-237.

Casellas,J., Tomas,A., Sanchez,A., Alves,E., Noguera,J.L., & Piedrafita,J. (2008). Using haplotype probabilities in categorical survival analysis: a case study with three candidate genes in an Iberian x Meishan F2 population of newborn piglets. *J Anim Breed Genet.*, 125(1), 5-12.

Cassady,J.P., Johnson,R.K., Pomp,D., Rohrer,G.A., Van Vleck,L.D., Spiegel,E.K., & Gilson,K.M. (2001). Identification of quantitative trait loci affecting reproduction in pigs. *J Anim Sci.*, 79(3), 623-633.

Chen,C., Chang,T., & Su,H. (2004a). Genetic polymorphisms in porcine leptin gene and their association with reproduction and production traits. *Australian.Journal of Agricultural Research*, 55(6), 699-704.

Chen,C.C., Chang,T., & Su,H.Y. (2004b). Characterization of porcine leptin receptor polymorphisms and their association with reproduction and production traits. *Anim Biotechnol.*, 15(1), 89-102.

Chen,C.Y., Guo,Y.M., Zhang,Z.Y., Ren,J., & Huang,L.S. (2010). A whole genome scan to detect quantitative trait loci for gestation length and sow maternal ability related traits in a White Duroc x Erhualian F_2 resource population. *animal*, 4(06), 861-866.

Chen,K.F., Huang,L.S., Li,N., Zhang,Q., Luo,M., & Wu,C.X. (2000). The genetic effect of estrogen receptor(ESR) on litter size traits in pig. *Yi Chuan Xue Bao.*, 27(10), 853-857.

Christenson,R.K., Leymaster,K.A., & Young,L.D. (1987). Justification of unilateral hysterectomy-ovariectomy as a model to evaluate uterine capacity in swine. *J Anim Sci.*, 65(3), 738-744.

Christenson,R.K., Vallet,J.L., Leymaster,K.A., & Young,L.D. (1993). Uterine function in Meishan pigs. *J Reprod Fertil Suppl.*, 48, 279-289.

Cintado,C.G., Pinto,F.M., Devillier,P., Merida,A., & Candenas,M.L. (2001). Increase in Neurokinin B Expression and in Tachykinin NK3 Receptor-Mediated Response and Expression in the Rat Uterus with Age. *J Pharmacol Exp Ther.*, 299(3), 934-938.

Cooper,D.K., Ezzelarab,M., Hara,H., & Ayares,D. (2008). Recent advances in pig-to-human organ and cell transplantation. *Expert Opin Biol Ther.*, 8(1), 1-4.

Coppieters,W., Van de Weghe,A., Peelman,L., VAN Zeveren,A., Bouquet,Y., & Depicker,A. (1993). Characterization of porcine polymorphic microsatellite loci. *Anim Genet.*, 24(3), 163-170.

Cox,M.L., Eddy,S.M., Stewart,Z.S., Kennel,M.R., Man,M.Z., Paulauskis,J.D., & Dunstan,R.W. (2008). Investigating fixative-induced changes in RNA quality and utility by microarray analysis. *Exp Mol Pathol.*, 84(2), 156-172.

Cox,M.L., Schray,C.L., Luster,C.N., Stewart,Z.S., Korytko,P.J., Khan,M., Paulauskis,J.D., & Dunstan,R.W. (2006). Assessment of fixatives, fixation, and tissue processing on morphology and RNA integrity. *Exp Mol Pathol.*, 80(2), 183-191.

Cunningham,P.J., England,M.E., Young,L.D., & Zimmerman,D.R. (1979). Selection for ovulation rate in swine: correlated response in litter size and weight. *J Anim Sci.*, 48(3), 509-516.

D'Souza,S.S., Daikoku,T., Farach-Carson,M.C., & Carson,D.D. (2007). Heparanase Expression and Function During Early Pregnancy in Mice. *Biol Reprod.*, 77(3), 433-441.

D'Souza,S.S., Fazleabas,A.T., Banerjee,P., Sherwin,J.R., Sharkey,A.M., Farach-Carson,M.C., & Carson,D.D. (2008). Decidual Heparanase Activity Is Increased During Pregnancy in the Baboon (Papio anubis) and in In Vitro Decidualization of Human Stromal Cells. *Biol Reprod.*, 78(2), 316-323.

Damgaard,L.H., Rydhmer,L., L+©vendahl,P., & Grandinson,K. (2003). Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. *Journal of Animal Science*, 81(3), 604-610.

Dantzer,V. & Leiser,R. (1994). Initial vascularisation in the pig placenta: I. Demonstration of nonglandular areas by histology and corrosion casts. *Anat Rec.*, 238(2), 177-190.

Dantzer,V. & Winther,H. (2001). Histological and immunohistochemical events during placentation in pigs. *Reprod Suppl.*, 58, 209-222.

Davis,G.P. & DeNise,S.K. (1998). The impact of genetic markers on selection. *J Anim Sci*, 76(9), 2331-2339.

Davis,K.L., Robison,O.W., & Toelle,V.D. (1987). Breed differences in uterine and ovarian measurements in gestating swine. *J Anim Sci.*, 65(3), 685-691.

de Koning,D.J., Rattink,A.P., Harlizius,B., Groenen,M.A.M., Brascamp,E.W., & van Arendonk,J.A.M. (2001). Detection and characterization of quantitative trait loci for growth and reproduction traits in pigs. *Liv Prod Sci.*, 72(3), 185-198.

de Vries,A.G., Sosnicki,A., Garnier,J.P., & Plastow,G.S. (1998). The role of major genes and DNA technology in selection for meat quality in pigs. *Meat.Sci*, 49S1, S245-S255.

Dempsey,L.A., Plummer,T.B., Coombes,S.L., & Platt,J.L. (2000). Heparanase expression in invasive trophoblasts and acute vascular damage. *Glycobiology*, 10(5), 467-475.

Denker,H.W. (1993). Implantation: a cell biological paradox. *J Exp Zool.*, 266(6), 541-558.

Ding,N., Guo,Y., Knorr,C., Ma,J., Mao,H., Lan,L., Xiao,S., Ai,H., Haley,C.S., Brenig,B., & Huang,L. (2009). Genome-wide QTL mapping for three traits related to teat number in a White Duroc x Erhualian pig resource population. *BMC Genet.*, 10, 6.

Drogemuller,C., Hamann,H., & Distl,O. (2001). Candidate gene markers for litter size in different German pig lines. *J Anim Sci.*, 79(10), 2565-2570.

Du,F.X., Clutter,A.C., & Lohuis,M.M. (2007). Characterizing linkage disequilibrium in pig populations. *Int J Biol Sci*, 3(3), 166-178.

Duijvesteijn,N., Knol,E.F., Merks,J.W., Crooijmans,R.P., Groenen,M.A., Bovenhuis,H., & Harlizius,B. (2010). A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genet.*, 11, 42.

Ellegren,H., Chowdhary,B.P., Johansson,M., Marklund,L., Fredholm,M., Gustavsson,I., & Andersson,L. (1994). A Primary Linkage Map of the Porcine Genome Reveals a Low Rate of Genetic Recombination. *Genetics*, 137(4), 1089-1100.

Ellegren,H., Johansson,M., Chowdhary,B.P., Marklund,S., Ruyter,D., Marklund,L., Brauner-Nielsen,P., Edfors-Lilja,I., Gustavsson,I., & Juneja,R.K. (1993). Assignment of 20 microsatellite markers to the porcine linkage map. *Genomics*, 16(2), 431-439.

Erikson,D.W., Burghardt,R.C., Bayless,K.J., & Johnson,G.A. (2009). Secreted Phosphoprotein 1 (SPP1, Osteopontin) Binds to Integrin Alphavbeta6 on Porcine Trophectoderm Cells and Integrin Alphavbeta3 on Uterine Luminal Epithelial Cells, and Promotes Trophectoderm Cell Adhesion and Migration. *Biol Reprod.*.

Estany,J., Villalba,D., Tibau,J., Soler,J., Babot,D., & Noguera,J.L. (2002). Correlated response to selection for litter size in pigs: I. Growth, fat deposition, and feeding behavior traits. *J Anim Sci.*, 80(10), 2556-2565.

Falconer,D.S. (1971). Improvement of litter size in a strain of mice at a selection limit. *Genet Res.*, 17(3), 215-235.

Falconer,D.S. & Mackay,T.F.C. (1996). *Introduction to Quantitative Genetics*. Harlow, Essex, UK: Addison Wesley Longman.

Fan,B., Onteru,S.K., Du,Z.Q., Garrick,D.J., Stalder,K.J., & Rothschild,M.F. (2011). Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs. *PLoS One.*, 6(2), e14726.

Fan, B., Onteru, S. K., Nikkila, M. T., Stalder, K. J., and Rothschild, M. F. Large-Scale Association Studies for Reproductive Traits in pigs. Plant and Animal genome XVII. Proceedings of the 17th Plant and Animal genome conference . 2009. San Diego, CA. Ref Type: Conference Proceeding

Fazleabas,A.T., Bell,S.C., Fleming,S., Sun,J., & Lessey,B.A. (1997). Distribution of integrins and the extracellular matrix proteins in the baboon endometrium during the menstrual cycle and early pregnancy. *Biol Reprod.*, 56(2), 348-356.

Fernandez,A., Rodrigañez,J., Zuzuarregui,J., Rodriguez,M.C., & Silio,L. (2008). Genetic parameters for litter size and weight at different parities in Iberian pigs. *Spanish Journal of Agricultural Research*, 6(Special issue), 00.

Fernandez-Rodriguez,A., Rodriguez,C., Varona,L., Balcells,I., Noguera,J.L., Ovilo,C., & Fernandez,A.I. (2010). Analysis of candidate genes underlying two epistatic quantitative trait loci on SSC12 affecting litter size in pig. *Anim Genet.*, 41(1), 73-80.

Fernandez-Rodriguez,A., Muñoz,M., Fernandez,A., Pena,R.N., Tomas,A., Noguera,J.L., Ovilo,C., & Fernandez,A.I. (2011). Differential Gene Expression in Ovaries of Pregnant Pigs with High and Low Prolificacy Levels and Identification of Candidate Genes for Litter Size. *Biol Reprod.*, 84(2), 299-307.

Finch,A.M., Antipatis,C., Pickard,A.R., & Ashworth,C.J. (2002). Patterns of fetal growth within Large White x Landrace and Chinese Meishan gilt litters at three stages of gestation. *Reprod Fertil Dev.*, 14(7-8), 419-425.

Fisher,L.W., Stubbs,J.T., III, & Young,M.F. (1995). Antisera and cDNA probes to human and certain animal model bone matrix noncollagenous proteins. *Acta Orthop Scand Suppl.*, 266, 61-65.

Florell,S.R., Coffin,C.M., Holden,J.A., Zimmermann,J.W., Gerwels,J.W., Summers,B.K., Jones,D.A., & Leachman,S.A. (2001). Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. *Mod.Pathol.*, 14(2), 116-128.

Ford,S.P. (1997). Embryonic and fetal development in different genotypes in pigs. *J Reprod Fertil Suppl.*, 52, 165-176.

Ford,S.P. & Youngs,C.R. (1993). Early embryonic development in prolific Meishan pigs. *J Reprod Fertil Suppl.*, 48, 271-278.

Fowler,D.J., Nicolaides,K.H., & Miell,J.P. (2000). Insulin-like growth factor binding protein-1 (IGFBP-1): a multifunctional role in the human female reproductive tract. *Hum Reprod Update.*, 6(5), 495-504.

Foxcroft,G.R., Bee,G., Dixon,W.T., Hahn,M.W., Harding,J.C., Patterson,J., Putman,T., Sarmento,S., Smit,M., Tse,W.-Y., & Town,S.C. (2007). Consequences of selection for liter size on piglet development. InJ.Wiseman, M.A.Varley, S.McOrist, & B.Kemp (Eds.), *Paradigms of Pig Science* (pp. 207-229). Nottingham, UK: Nottingham University Press.

Foxcroft,G.R., Dixon,W.T., Dyck,M.K., Novak,S., Harding,J.C., & Almeida,F.C. (2009). Prenatal programming of postnatal development in the pig. *Soc Reprod Fertil Suppl.*, 66, 213-231.

Foxcroft,G.R., Dixon,W.T., Novak,S., Putman,C.T., Town,S.C., & Vinsky,M.D. (2006). The biological basis for prenatal programming of postnatal performance in pigs. *J Anim Sci.*, 84 Suppl, E105-E112.

Freeman,C. & Parish,C.R. (1998). Human platelet heparanase: purification, characterization and catalytic activity. *Biochem J.*, 330 (Pt 3), 1341-1350.

Friess,A.E., Sinowatz,F., Skolek-Winnisch,R., & Traautner,W. (1980). The placenta of the pig. I. Finestructural changes of the placental barrier during pregnancy. *Anat.Embryol.(Berl)*, 158(2), 179-191.

Galvin,J.M., Wilmut,I., Day,B.N., Ritchie,M., Thomson,M., & Haley,C.S. (1993). Reproductive performance in relation to uterine and embryonic traits during early gestation in Meishan, large white and crossbred sows. *J Reprod Fertil.*, 98(2), 377-384.

Garlow,J.E., Ka,H., Johnson,G.A., Burghardt,R.C., Jaeger,L.A., & Bazer,F.W. (2002). Analysis of osteopontin at the maternal-placental interface in pigs. *Biol Reprod.*, 66(3), 718-725.

Geisert,R.D., Renegar,R.H., Thatcher,W.W., Roberts,R.M., & Bazer,F.W. (1982). Establishment of pregnancy in the pig: I. Interrelationships between preimplantation development of the pig blastocyst and uterine endometrial secretions. *Biol Reprod.*, 27(4), 925-939.

Geisert,R.D. & Schmitt,R.A.M. (2002). Early embryonic survival in the pig: Can it be improved? *J Anim Sci.*, 80(E-Suppl_1), E54-E65.

Geisert,R.D., zavy,M.T., Moffatt,R.J., Blair,R.M., & Yellin,T. (1990). Embryonic steroids and the establishment of pregnancy in pigs. *J Reprod Fertil Suppl.*, 40, 293-305.

Giancotti,F.G. & Ruoslahti,E. (1999). Integrin signaling. *Science*, 285(5430), 1028-1032.

Gibson,J.P., Jiang,Z.H., Robinson,J.A., Archibald,A.L., & Haley,C.S. (2002). No detectable association of the ESR PvuII mutation with sow productivity in a Meishan x Large White F2 population. *Anim Genet.*, 33(6), 448-450.

Gilmour, A. R, Gogel, B. J., Cullis, B. R., and Thompson, R. ASReml User Guide Release 3.0. 2009. Hempstead, UK, VSN International Ltd.
Ref Type: Computer Program

Gladney,C.D., Bertani,G.R., Johnson,R.K., & Pomp,D. (2004). Evaluation of gene expression in pigs selected for enhanced reproduction using differential display PCR and human microarrays: I. Ovarian follicles. *J Anim Sci.*, 82(1), 17-31.

Gladney,C.D., Martinez,V.G., Brumbaugh,K.A., DeGroot,B.J., Linville,R.C., Oommen,A.M., Huebinger,R.M., Messer,L.A., Allan,M.F., & Pomp,D. (1999). Mapping of the Prostaglandin-Endoperoxide Synthase 2 (PTGS2) gene to porcine chromosome 9 and bovine chromosome 16 by linkage analysis using novel PCR-RFLP. *J Anim Sci.*, 77(3), 787-788.

Goliasova,E. & Wolf,J. (2004). Impact of the ESR gene on litter size and production traits in Czech Large White pigs. *Anim Genet.*, 35(4), 293-297.

Goshen,R., Hochberg,A.A., Korner,G., Levy,E., Ishai-Michaeli,R., Elkin,M., de,G.N., & Vlodavsky,I. (1996). Purification and characterization of placental heparanase and its expression by cultured cytotrophoblasts. *Mol Hum Reprod.*, 2(9), 679-684.

Green, P., Falls, K., and Crooks, S. Documentation for CRI-MAP, Version 2.4. 1990. St.Louis, Mo., Washington Univesity School of Medicine.
Ref Type: Computer Program

Grindflek,E., Meuwissen,T.H.E., Aasmundstad,T., Hamland,H., Hansen,M.H.S., Nome,T., Kent,M., Torjesen,P., & Lien,S. (2011). Revealing genetic relationships between compounds affecting boar taint and reproduction in pigs. *J Anim Sci.*, 89(3), 680-692.

Guo,Y.M., Lee,G.J., Archibald,A.L., & Haley,C.S. (2008). Quantitative trait loci for production traits in pigs: a combined analysis of two Meishan x Large White populations. *Animal Genetics*, 39(5), 486-495.

Guo,Y.M., Mao,H., Ren,J., Yan,X., Duan,Y., Yang,G., Ren,D., Zhang,Z.Y., Yang,B., Ouyang,J., Brenig,B., Haley,C.S., & Huang,L. (2009). A linkage map of the porcine genome from a large-scale White Duroc x Erhualian resource population and evaluation of factors affecting recombination rates. *Anim Genet.*, 40(1), 47-52.

Haimov-Kochman,R., Friedmann,Y., Prus,D., Goldman-Wohl,D.S., Greenfield,C., Anteby,E.Y., Aviv,A., Vlodavsky,I., & Yagel,S. (2002). Localization of heparanase in normal and pathological human placenta. *Mol Hum Rep.*, 8(6), 566-573.

Haley,C.S., d'Agaro,E., & Ellis,M. (1992). Genetic components of growth and ultrasonic fat depth traits in Meishan and Large White pigs and their reciprocal crosses. *Anim Sci.*, 54(01), 105-115.

Haley,C.S., Knott,S.A., & Elsen,J.M. (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics*, 136(3), 1195-1207.

Haley, C. S. and Lee, G. J. Genetic components of litter size in Meishan and Large White pigs and their crosses. 4th World Congress on Genetics Applied to Livestock Production XV. Proceedings of the 4th World Congress on Genetics Applied to Livestock Production XV , 458-481. 1990.
Ref Type: Conference Proceeding

Haley,C.S. & Lee,G.J. (1993). Genetic basis of prolificacy in Meishan pigs. *J Reprod Fertil Suppl.*, 48, 247-259.

Haley,C.S., Lee,G.J., & Ritchie,M. (1995). Comparative reproductive performance in Meishan and Large White Pigs and their crosses. *Animal Science*, 60, 259-267.

Hamatani,K., Eguchi,H., Takahashi,K., Koyama,K., Mukai,M., Ito,R., Taga,M., Yasui,W., & Nakachi,K. (2006). Improved RT-PCR amplification for molecular analyses with long-term preserved formalin-fixed, paraffin-embedded tissue specimens. *Journal of Histochemistry and Cytochemistry*, 54(7), 773-780.

Hanenberg,E.H.A.T., Knol,E.F., & Merks,J.W.M. (2001). Estimates of genetic parameters for reproduction traits at different parities in Dutch Landrace pigs. *Liv Prod Sci.*, 69(2), 179-186.

Hao,Y., Murphy,C.N., Spate,L., Wax,D., Zhong,Z., Samuel,M., Mathialagan,N., Schatten,H., & Prather,R.S. (2008). Osteopontin improves in vitro development of porcine embryos and decreases apoptosis. *Mol Reprod Dev.*, 75(2), 291-298.

Harney,J.P., Mirando,M.A., Smith,L.C., & Bazer,F.W. (1990). Retinol-binding protein: a major secretory product of the pig conceptus. *Biol Reprod.*, 42(3), 523-532.

Hasengaowa, Kodama,J., Kusumoto,T., Seki,N., Matsuo,T., Ojima,Y., Nakamura,K., Hongo,A., & Hiramatsu,Y. (2006). Heparanase expression in both normal endometrium and endometrial cancer. *Int J Gynecol Cancer.*, 16(3), 1401-1406.

Hellbrugge,B., Tolle,K.H., Bennewitz,J., Henze,C., Presuhn,U., & Krieter,J. (2008). Genetic aspects regarding piglet losses and the maternal behaviour of sows. Part 1. Genetic analysis of piglet mortality and fertility traits in pigs. *animal*, 2(9), 1273-1280.

Hellemans,J., Mortier,G., De,P.A., Speleman,F., & Vandesompele,J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.*, 8(2), R19.

Herington,J.L. & Bany,B.M. (2007). The conceptus increases secreted phosphoprotein 1 gene expression in the mouse uterus during the progression of decidualization mainly due to its effects on uterine natural killer cells. *Reprod.*, 133(6), 1213-1221.

Herpin,P., Damon,M., & Le Dividich,J. (2002). Development of thermoregulation and neonatal survival in pigs. *Liv Prod Sci.*, 78(1), 25-45.

Hirooka,H., de Koning,D.J., Harlizius,B., van Arendonk,J.A., Rattink,A.P., Groenen,M.A., Brascamp,E.W., & Bovenhuis,H. (2001). A whole-genome scan for quantitative trait loci affecting teat number in pigs. *J Anim Sci.*, 79(9), 2320-2326.

Holl,J.W., Cassady,J.P., Pomp,D., & Johnson,R.K. (2004). A genome scan for quantitative trait loci and imprinted regions affecting reproduction in pigs. *J Anim Sci.*, 82(12), 3421-3429.

Horak,P., Urban,T., & Dvorak,J. (2005). The FUT1 and ESR genes--their variability and associations with reproduction in Prestice Black-Pied sows. *J Anim Breed.Genet*, 122(3), 210-213.

Horogh,G., Zsolnai,A., Komiosi,I., Nyiri,A., Anton,I., & Fesus,L. (2005). Oestrogen receptor genotypes and litter size in Hungarian Large White pigs. *J Anim Breed Genet.*, 122(1), 56-61.

Hume,D.A., Whitelaw,C.B.A., & Archibald,A.L. (2011). The future of animal production: improving productivity and sustainability. *The Journal of Agricultural Science*, 149(Supplement S1), 9-16.

Humphray,S.J., Scott,C.E., Clark,R., Marron,B., Bender,C., Camm,N., Davis,J., Jenks,A., Noon,A., Patel,M., Sehra,H., Yang,F., Rogatcheva,M.B., Milan,D., Chardon,P., Rohrer,G., Nonneman,D., de,J.P., Meyers,S.N., Archibald,A., Beever,J.E., Schook,L.B., & Rogers,J. (2007). A high utility integrated map of the pig genome. *Genome Biol.*, 8(7), R139.

Irgang,R., Favero,J.A., & Kennedy,B.W. (1994). Genetic parameters for litter size of different parities in Duroc, Landrace, and large white sows. *J Anim Sci.*, 72(9), 2237-2246.

Isler,B.J., Irvin,K.M., Neal,S.M., Moeller,S.J., & Davis,M.E. (2002). Examination of the relationship between the estrogen receptor gene and reproductive traits in swine. *J Anim Sci.*, 80(9), 2334-2339.

Jiang,Z., Gibson,J.P., Archibald,A.L., & Haley,C.S. (2001). The porcine gonadotropin-releasing hormone receptor gene (GNRHR): genomic organization, polymorphisms, and association with the number of corpora lutea. *Genome*, 44(1), 7-12.

Jiang,Z., He,H., Hamasima,N., Suzuki,H., & Verrinder,G. (2002a). Comparative mapping of Homo sapiens chromosome 4 (HSA4) and Sus scrofa chromosome 8 (SSC8) using orthologous genes representing different cytogenetic bands as landmarks. *Genome*, 45(1), 147-156.

Jiang, Z., Robinson, J. A., Verrinder, A. M., Gibson, J. P., Archibald, A. L., and Haley, C. S. Mapping of QTLs for prolificacy traits on SSC8 using a candidate gene approach. 7th World Congress on Genetics Applied to Livestock Production Session 08. Reproduction. 2002b. Ref Type: Conference Proceeding

Johnson,G.A., Bazer,F.W., Burghardt,R.C., Spencer,T.E., Wu,G., & Bayless,K.J. (2009). Conceptus-uterus interactions in pigs: endometrial gene expression in response to estrogens and interferons from conceptuses. *Soc Reprod Fertil Suppl.*, 66, 321-332.

Johnson,G.A., Burghardt,R.C., Bazer,F.W., & Spencer,T.E. (2003a). Osteopontin: roles in implantation and placentation. *Biol Reprod.*, 69(5), 1458-1471.

Johnson,G.A., Burghardt,R.C., Joyce,M.M., Spencer,T.E., Bazer,F.W., Gray,C.A., & Pfarrer,C. (2003b). Osteopontin is synthesized by uterine glands and a 45-kDa cleavage fragment is localized at the uterine-placental interface throughout ovine pregnancy. *Biol Reprod.*, 69(1), 92-98.

Johnson,G.A., Burghardt,R.C., Joyce,M.M., Spencer,T.E., Bazer,F.W., Pfarrer,C., & Gray,C.A. (2003c). Osteopontin expression in uterine stroma indicates a decidualization-like differentiation during ovine pregnancy. *Biol Reprod.*, 68(6), 1951-1958.

Johnson,G.A., Burghardt,R.C., Spencer,T.E., Newton,G.R., Ott,T.L., & Bazer,F.W. (1999a). Ovine osteopontin: II. Osteopontin and alpha(v)beta(3) integrin expression in the uterus and conceptus during the periimplantation period. *Biol Reprod.*, 61(4), 892-899.

Johnson,G.A., Spencer,T.E., Burghardt,R.C., & Bazer,F.W. (1999b). Ovine osteopontin: I. Cloning and expression of messenger ribonucleic acid in the uterus during the periimplantation period. *Biol Reprod.*, 61(4), 884-891.

Johnson,G.A., Spencer,T.E., Hansen,T.R., Austin,K.J., Burghardt,R.C., & Bazer,F.W. (1999c). Expression of the interferon tau inducible ubiquitin cross-reactive protein in the ovine uterus. *Biol Reprod.*, 61(1), 312-318.

Johnson,R.K., Nielsen,M.K., & Casey,D.S. (1999d). Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. *J Anim Sci.*, 77(3), 541-557.

Jonas,E., Schreinemachers,H.J., Kleinwachter,T., Un,C., Oltmanns,I., Tetzlaff,S., Jennen,D., Tesfaye,D., Ponsuksili,S., Murani,E., Juengst,H., Tholen,E., Schellander,K., & Wimmers,K. (2008). QTL for the heritable inverted teat defect in pigs. *Mamm Genome.*, 19(2), 127-138.

Joyce,M.M., Gonzalez,J.F., Lewis,S., Woldesenbet,S., Burghardt,R.C., Newton,G.R., & Johnson,G.A. (2005). Caprine uterine and placental osteopontin expression is distinct among epitheliochorial implanting species. *Placenta*, 26(2-3), 160-170.

Ka,H., Seo,H., Kim,M., Choi,Y., & Lee,C.K. (2009). Identification of differentially expressed genes in the uterine endometrium on day 12 of the estrous cycle and pregnancy in pigs. *Mol Reprod Dev.*, 76(1), 75-84.

Kao,L.C., Tulac,S., Lobo,S., Imani,B., Yang,J.P., Germeyer,A., Osteen,K., Taylor,R.N., Lessey,B.A., & Giudice,L.C. (2002). Global gene profiling in human endometrium during the window of implantation. *Endocrinology*, 143(6), 2119-2138.

Kearsey,M.J. (1998). The principles of QTL analysis (a minimal mathematics approach). *J Exp Bot.*, 49(327), 1619.

Keats,B.J.B., Sherman,S.L., Morton,N.E., Robson,E.B., Buetow,K.H., Cartwright,P.E., Chakravarti,A., Francke,U., Green,P.P., & Ott,J. (1991). Guidelines for human linkage maps: An international system for human linkage maps (ISLM, 1990). *Genomics*, 9(3), 557-560.

Kim,J.G., Song,J.H., Vallet,J.L., Rohrer,G.A., Johnson,G.A., Joyce,M.M., & Christenson,R.K. (2003). Molecular characterization and expression of porcine bone morphogenetic protein receptor-IB in the uterus of cyclic and pregnant gilts. *Biol Reprod.*, 68(3), 735-743.

Kim,J.G., Vallet,J.L., Rohrer,G.A., & Christenson,R.K. (2002). Characterization of porcine uterine estrogen sulfotransferase. *Domest Anim Endocrinol*, 23(4), 493-506.

Kim,J., Erikson,D.W., Burghardt,R.C., Spencer,T.E., Wu,G., Bayless,K.J., Johnson,G.A., & Bazer,F.W. (2010). Secreted phosphoprotein 1 binds integrins to initiate multiple cell

signaling pathways, including FRAP1/mTOR, to support attachment and force-generated migration of trophectoderm cells. *Matrix Biol.*, 29(5), 369-382.

King,A.H. (2003). Identification of Candidate Genes Controlling Porcine Female Reproductive Traits.

King,A.H., Haley,C.S., & Archibald,A.L. (2001). Quantitative trait loci (QTL) for female reproductive traits on porcine chromosome 8. *Genet Res.*, 78(2), 191-206.

King,A.H., Jiang,Z., Gibson,J.P., Haley,C.S., & Archibald,A.L. (2003). Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8. *Biol Reprod.*, 68(6), 2172-2179.

Kizaki,K., Nakano,H., Nakano,H., Takahashi,T., Imai,K., & Hashizume,K. (2001). Expression of heparanase mRNA in bovine placenta during gestation. *Reprod.*, 121(4), 573-580.

Knoll,A., Stratil,A., Cepica,S., & Dvorak,J. (1999). Length polymorphism in an intron of the porcine osteopontin (SPP1) gene is caused by the presence or absence of a SINE (PRE-1) element. *Anim Genet.*, 30(6), 466.

Knott,S.A. & Haley,C.S. (2000). Multitrait Least Squares for Quantitative Trait Loci Detection. *Genetics*, 156(2), 899-911.

Korwin-Kossakowska,A., Kamyczek,M., Cieslak,D., & Kuryl,J. (2001). The polymorphism of reproduction-linked genes in Line 900 sows. *Animal Science Papers and Reports.*, 19(4), 265-274.

Korwin-Kossakowska,A., Kamyczek,M., Cieslak,D., Pierzchala,M., & Kuryl,J. (2002). The effect of the polymorphism of leptin (LEP), leptin receptor (LEPR) and osteopontin (OPN) genes on selected reproduction traits of synthetic Line 990 sow. *Animal Science Papers and Reports*, 20(3), 159-168.

Kuehn,L.A., Rohrer,G.A., Nonneman,D.J., Thallman,R.M., & Leymaster,K.A. (2007). Detection of single nucleotide polymorphisms associated with ultrasonic backfat depth in a segregating Meishan x White Composite population. *J Anim Sci.*, 85(5), 1111-1119.

Kuijk,E.W., du,P.L., van Tol,H.T., Haagsman,H.P., Colenbrander,B., & Roelen,B.A. (2007). Validation of reference genes for quantitative RT-PCR studies in porcine oocytes and preimplantation embryos. *BMC Dev Biol.*, 7, 58.

Lamberson,W.R., Johnson,R.K., Zimmerman,D.R., & Long,T.E. (1991). Direct responses to selection for increased litter size, decreased age at puberty, or random selection following selection for ovulation rate in swine. *J Anim Sci.*, 69(8), 3129-3143.

Lan,J., Zhao,J., & Liu,Y. (2011). Molecular cloning, sequence characterization, polymorphism and association analysis of porcine ROPN1 gene. *Mol Hum Rep.*, 1-5.

Lande,R. & Thompson,R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 124(3), 743-756.

Lee,K.Y., Shibutani,M., Inoue,K., Kuroiwa,K., U,M., Woo,G.H., & Hirose,M. (2006). Methacarn fixation--effects of tissue processing and storage conditions on detection of mRNAs and proteins in paraffin-embedded tissues. *Analytical Biochemistry*, 351(1), 36-43.

Lee,S.S., Chen,Y., Moran,C., Stratil,A., Reiner,G., Bartenschlager,H., Moser,G., & Geldermann,H. (2003). Linkage and QTL mapping for Sus scrofa chromosome 5. *J Anim Breed Genet.*, 120, 38-44.

Leenhouwers,J.I., Knol,E.F., de Groot,P.N., Vos,H., & van der Lende,T. (2002). Fetal development in the pig in relation to genetic merit for piglet survival. *J Anim Sci.*, 80(7), 1759-1770.

Leiser,R. & Dantzer,V. (1988). Structural and functional aspects of porcine placental microvasculature. *Anat Embryol.(Berl)*, 177(5), 409-419.

Lessey,B.A.P. (1995). Endometrial Integrins. [Editorial]. *Endocrinologist*, 5(3), 214-221.

Lewis,C.R.G., Torremorell,M., & Bishop,S.C. (2009a). Effects of porcine reproductive and respiratory syndrome virus infection on the performance of commercial sows and gilts of different parities and genetic lines. *J Swine Health Prod.*, 17(3), 140-147.

Lewis,C.R.G., Torremorell,M., Galina-Pantoja,L., & Bishop,S.C. (2009b). Genetic parameters for performance traits in commercial sows estimated before and after an outbreak of porcine reproductive and respiratory syndrome. *J Anim Sci.*, 87(3), 876-884.

Li,K., Ren,J., Xing,Y., Zhang,Z., Ma,J., Guo,Y., & Huang,L. (2009). Quantitative trait loci for litter size and prenatal loss in a White Duroc x Chinese Erhualian resource population. *Anim Genet.*.

Li,M.D., Macdonald,G.J., Wise,T., & Ford,J.J. (1998). Positive association between expression of follicle-stimulating hormone beta and activin betaB-subunit genes in boars. *Biol Reprod.*, 59(4), 978-982.

Li,X.P., Hu,Z.L., Moon,S.J., Do,K.T., Ha,Y.K., Kim,H., Byun,M.J., Choi,B.H., Rothschild,M.F., Reecy,J.M., & Kim,K.S. (2008). Development of an in silico coding gene SNP map in pigs. *Anim Genet.*, 39(4), 446-450.

Lin,H.C., Liu,G.F., Wang,A.G., Kong,L.J., Wang,X.F., & Fu,J.L. (2009). Effect of polymorphism in the leukemia inhibitory factor gene on litter size in Large White pigs. *Mol Biol Rep.*, 36(7), 1833-1838.

Linville,R.C., Pomp,D., Johnson,R.K., & Rothschild,M.F. (2001). Candidate gene analysis for loci affecting litter size and ovulation rate in swine. *J Anim Sci.*, 79(1), 60-67.

Lo Buono,N., Parrotta,R., Morone,S., Bovino,P., Nacci,G., Ortolan,E., Horenstein,A.L., Inzhutova,A., Ferrero,E., & Funaro,A. (2011). The CD157-integrin partnership controls transendothelial migration and adhesion of human monocytes. *J Biol Chem.*, 286(21), 18681-18691.

Marklund,L., Johansson,M.M., Hoyheim,B., Davies,W., Fredholm,M., Juneja,R.K., Mariani,P., Coppieters,W., Ellegren,H., & Andersson,L. (1996). A comprehensive linkage map of the pig based on a wild pig-Large White intercross. *Anim Genet.*, 27(4), 255-269.

Martinez-Giner,M., Noguera,J.L., Balcells,I., Alves,E., Varona,L., & Pena,R.N. (2011). Expression study on the porcine PTHLH gene and its relationship with sow teat number. *J Anim Breed Genet.*, 128(5), 344-353.

Matise,T.C., Perlin,M., & Chakravarti,A. (1994). Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. *Nat Genet.*, 6(4), 384-390.

McNeil,C.J., Nwagwu,M.O., Finch,A.M., Page,K.R., Thain,A., McArdle,H.J., & Ashworth,C.J. (2007). Glucocorticoid exposure and tissue gene expression of 11beta HSD-1, 11beta HSD-2, and glucocorticoid receptor in a porcine model of differential fetal growth. *Reproduction*, 133(3), 653-661.

McPherson,J.D., Marra,M., Hillier,L., Waterston,R.H., Chinwalla,A., Wallis,J., Sekhon,M., Wylie,K., Mardis,E.R., Wilson,R.K., Fulton,R., Kucaba,T.A., Wagner-McPherson,C., Barbazuk,W.B., Gregory,S.G., Humphray,S.J., French,L., Evans,R.S., Bethel,G., Whittaker,A., Holden,J.L., McCann,O.T., Dunham,A., Soderlund,C., Scott,C.E., Bentley,D.R., Schuler,G., Chen,H.C., Jang,W., Green,E.D., Idol,J.R., Maduro,V.V., Montgomery,K.T., Lee,E., Miller,A., Emerling,S., Kucherlapati, Gibbs,R., Scherer,S., Gorrell,J.H., Sodergren,E., Clerc-Blankenburg,K., Tabor,P., Naylor,S., Garcia,D., de Jong,P.J., Catanese,J.J., Nowak,N., Osoegawa,K., Qin,S., Rowen,L., Madan,A., Dors,M., Hood,L., Trask,B., Friedman,C., Massa,H., Cheung,V.G., Kirsch,I.R., Reid,T., Yonescu,R., Weissenbach,J., Bruls,T., Heilig,R., Branscomb,E., Olsen,A., Doggett,N., Cheng,J.F., Hawkins,T., Myers,R.M., Shang,J., Ramirez,L., Schmutz,J., Velasquez,O., Dixon,K., Stone,N.E., Cox,D.R., Haussler,D., Kent,W.J., Furey,T., Rogic,S., Kennedy,S., Jones,S., Rosenthal,A., Wen,G., Schilhabel,M., Gloeckner,G., Nyakatura,G., Siebert,R., Schlegelberger,B., Korenberg,J., Chen,X.N., Fujiyama,A., Hattori,M., Toyoda,A., Yada,T., Park,H.S., Sakaki,Y., Shimizu,N., Asakawa,S., Kawasaki,K., Sasaki,T., Shintani,A., Shimizu,A., Shibuya,K., Kudoh,J., Minoshima,S., Ramser,J., Seranski,P., Hoff,C., Poustka,A., Reinhardt,R., & Lehrach,H. (2001). A physical map of the human genome. *Nature*, 409(6822), 934-941.

Meller,M., Vadachkoria,S., Luthy,D.A., & Williams,M.A. (2005). Evaluation of housekeeping genes in placental comparative expression studies. *Placenta*, 26(8-9), 601-607.

Melville, J. S., Gibbins, A. M., Robinson, J. A., Gibson, J. P., Archibald, A. L., Haley, C. S., and Jiang, Z. A Meishan positive QTL for prolificacy traits found at the *NCOA1* locus on SSC3. Organising Committee Castanet-Tolosan. Proceedings of the 7th World Congress on Genetics Applied to Livestock Production 30, 733-736. 2002. France, INRA. Ref Type: Conference Proceeding

Mendez,E.A., Messer,L.A., Larsen,N.J., Robic,A., & Rothschild,M.F. (1999). Epidermal growth factor maps to pig chromosome 8. *J Anim Sci.*, 77(2), 494-495.

Mesa,H., Safranski,T.J., Fischer,K.A., Cammack,K.M., & Lamberson,W.R. (2005). Selection for placental efficiency in swine: genetic parameters and trends. *J Anim Sci.*, 83(5), 983-991.

Mesa,H., Safranski,T.J., Johnson,R.K., & Lamberson,W.R. (2003). Correlated response in placental efficiency in swine selected for an index of components of lifter size. *J Anim Sci.*, 81(1), 74-79.

Meuwissen,T.H., Hayes,B.J., & Goddard,M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819-1829.

Meyers,S.N., Rogatcheva,M.B., Larkin,D.M., Yerle,M., Milan,D., Hawken,R.J., Schook,L.B., & Beever,J.E. (2005). Piggy-BACing the human genome II. A high-resolution, physically anchored, comparative map of the porcine autosomes. *Genomics*, 86(6), 739-752.

Miese-Looy,G., VAN DEN Heuvel,M.J., Edwards,A.K., Lamarre,J., & Tayade,C. (2012). Expression of insulin-like growth factor (IGF) family members in porcine pregnancy. *J Reprod Dev.*, 58(1), 51-60.

Mikawa,S., Akita,T., Hisamatsu,N., Inage,Y., Ito,Y., Kobayashi,E., Kusumoto,H., Matsumoto,T., Mikami,H., Minezawa,M., Miyake,M., Shimanuki,S., Sugiyama,C., Uchida,Y., Wada,Y., Yanai,S., & Yasue,H. (1999). A linkage map of 243 DNA markers in an intercross of Gottingen miniature and Meishan pigs. *Anim Genet.*, 30(6), 407-417.

Miles,J.R., Vallet,J.L., Freking,B.A., & Nonneman,D.J. (2009). Molecular cloning and characterisation of heparanase mRNA in the porcine placenta throughout gestation. *Reprod Fertil Dev.*, 21(6), 757-772.

Mitchell,D., Ibrahim,S., & Gusterson,B.A. (1985). Improved immunohistochemical localization of tissue antigens using modified methacarn fixation. *J Histochem Cytochem.*, 33(5), 491-495.

Montgomery,G.W., Lord,E.A., Penty,J.M., Dodds,K.G., Broad,T.E., Cambridge,L., Sunden,S.L., Stone,R.T., & Crawford,A.M. (1994). The Booroola fecundity (FecB) gene maps to sheep chromosome 6. *Genomics*, 22(1), 148-153.

Moran,C. (1993). Microsatellite Repeats in Pig (Sus domestica) and Chicken (Gallus domesticus) Genomes. *J Hered.*, 84(4), 274-280.

Muñoz,A., Fernandez,A.I., Ovilo,C., Muñoz,G., Rodriguez,C., Fernandez,A., Alves,E., & Silio,L. (2010). Non-additive effects of RBP4, ESR1 and IGF2 polymorphisms on litter size at different parities in a Chinese-European porcine line. *Genet Sel Evol.*, 42(1), 23.

Muñoz,G., Ovilo,C., Amills,M., & Rodriguez,C. (2004). Mapping of the porcine oestrogen receptor 2 gene and association study with litter size in Iberian pigs. *Anim Genet.*, 35(3), 242-244.

Nadeau,J.H. & Sankoff,D. (1998). Counting on comparative maps. *Trends Genet.*, 14(12), 495-501.

Niu,B.Y., Ye,L.Z., Li,F.E., Deng,C.Y., Jiang,S.W., Lei,M.G., & Xiong,Y.Z. (2009). Identification of polymorphism and association analysis with reproductive traits in the porcine RNF4 gene. *Anim Reprod Sci*, 110(3-4), 283-292.

Noguera,J.L., Rodriguez,C., Varona,L., Tomas,A., Muñoz,G., Ramirez,O., Barragan,C., Arque,M., Bidanel,J.P., Amills,M., Ovilo,C., & Sanchez,A. (2009). A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. *BMC Genomics*, 10, 636.

Noguera, J. L., Rodriguez, M. C., Varona, L., Tomas, A., Muñoz, G., Ramirez, O., Barragan, C., Arque, M., Bidanel, J. P., Amills, M., Ovilo, C., and Sanchez, A. Epistasis is a

fundamental component of the genetic architecture of prolificacy in pigs. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production . 2006. 13-8-2006. Ref Type: Conference Proceeding

Noguera,J.L., Varona,L., Babot,D., & Estany,J. (2002). Multivariate analysis of litter size for multiple parities with production traits in pigs: II. Response to selection for litter size and correlated response to production traits. *J Anim Sci*, 80(10), 2548-2555.

Nomura,S., Wills,A.J., Edwards,D.R., Heath,J.K., & Hogan,B.L. (1988). Developmental expression of 2ar (osteopontin) and SPARC (osteonectin) RNA as revealed by in situ hybridization. *J Cell Biol.*, 106(2), 441-450.

Nonneman,D.J. & Rohrer,G.A. (2003). Comparative mapping of a region on chromosome 10 containing QTL for reproduction in swine. *Anim Genet.*, 34(1), 42-46.

Norwitz,E.R., Schust,D.J., & Fisher,S.J. (2001). Implantation and the Survival of Early Pregnancy. *N Engl J Med.*, 345(19), 1400-1408.

Nygard,A.B., Jorgensen,C.B., Cirera,S., & Fredholm,M. (2007). Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Mol Biol.*, 8, 67.

Onteru,S.K., Fan,B., Du,Z.Q., Garrick,D.J., Stalder,K.J., & Rothschild,M.F. (2012). A whole-genome association study for pig reproductive traits. *Animal Genetics*, 43(1), 18-26.

Onteru,S.K., Fan,B., Nikkila,M.T., Garrick,D.J., Stalder,K.J., & Rothschild,M.F. (2011). Whole-genome association analyses for lifetime reproductive traits in the pig. *J Anim Sci.*, 89(4), 988-995.

Onteru,S.K., Ross,J.W., & Rothschild,M.F. (2009). The role of gene discovery, QTL analyses and gene expression in reproductive traits in the pig. *Soc Reprod Fertil Suppl.*, 66, 87-102.

Orr Jr, D. E. and Shen, Y. World Pig Production, Opportunity or Threat? Swine Nutrition conference. Swine Nutrition Conference Proceeding , 3-8. 2006. Ref Type: Conference Proceeding

Page,N.M., Dakour,J., & Morrish,D.W. (2006). Gene regulation of neurokinin B and its receptor NK3 in late pregnancy and pre-eclampsia. *Mol Hum Rep.*, 12(7), 427-433.

Page,N.M., Woods,R.J., Gardiner,S.M., Lomthaisong,K., Gladwell,R.T., Butlin,D.J., Manyonda,I.T., & Lowry,P.J. (2000). Excessive placental secretion of neurokinin B during the third trimester causes pre-eclampsia. *Nature*, 405(6788), 797-800.

Page,N. (2010). Neurokinin B and pre-eclampsia: a decade of discovery. *Reprod Biol Endocrinol.*, 8(1), 4.

Perry,J.S. & Rowell,J.G. (1969). Variations in foetal weight and vascular supply along the uterine horn of the pig. *J Reprod Fertil.*, 19(3), 527-534.

Pickard,A.R. & Ashworth,C.J. (1995). Oestradiol-17β secretion by Chinese Meishan and Large White x Landrace conceptuses. *J Reprod Fertil Abstr Ser.*, 15, 49.

Pickard,A.R., Miller,S.J., & Ashworth,C.J. (2003). Synchronous onset of oestradiol-17beta secretion by Meishan conceptuses. *Reprod Biol Endocrinol.*, 1, 16.

Pinto,F.M., Cintado,C.G., Devillier,P., & Candenas,M.L. (2001). Expression of preprotachykinin-B, the gene that encodes neurokinin B, in the rat uterus. *Eur J Pharmacol.*, 425(2), R1-R2.

Pinto,F.M., Magraner,J., Ausina,P., Anselmi,E., Martin,J.D., & Candenas,M.L. (1997). Regulation by oestrogens of tachykinin NK3 receptor expression in the rat uterus. *Eur J Pharmacol.*, 324(1), 125-127.

Pinto,F.M., Pintado,C.O., Pennefather,J.N., Patak,E., & Candenas,M.L. (2009). Ovarian steroids regulate tachykinin and tachykinin receptor gene expression in the mouse uterus. *Reprod Biol Endocrinol.*, 7(1), 77.

Polge,C., Rowson,L.E., & Chang,M.C. (1966). The effect of reducing the number of embryos during early stages of gestation on the maintenance of pregnancy in the pig. *J Reprod Fertil*, 12(2), 395-397.

Pomp,D., Caetano,A.R., Bertani,G.R., Gladney,C.D., & Johnson,R.K. (2001). Applying functional genomics research to the study of pig reproduction. *Reprod Suppl.*, 58, 277-292.

Ponsuksili,S., Murani,E., Brand,B., Schwerin,M., & Wimmers,K. (2011). Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. *J Lipid Res.*, 52(4), 668-678.

Pope,W.F. (1994). Embryonic mortality in swine. InR.D.Geisert & M.T.zavy (Eds.), *Embryonic Mortality in Domestic Species* (pp. 53-78). London: Boca Raton.

Pope,W.F., Maurer,R.R., & Stormshak,F. (1982). Survival of porcine embryos after asynchronous transfer (41495). *Proc Soc Exp Biol Med.*, 171(2), 179-183.

Pope,W.F., Xie,S., Broermann,D.M., & Nephew,K.P. (1990). Causes and consequences of early embryonic diversity in pigs. *J Reprod Fertil Suppl.*, 40, 251-260.

Price,A.L., Patterson,N.J., Plenge,R.M., Weinblatt,M.E., Shadick,N.A., & Reich,D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.*, 38(8), 904-909.

Prunier,A., Heinonen,M., & Quesnel,H. (2010). High physiological demands in intensively raised pigs: impact on health and welfare. *animal*, 4(6), 886-898.

Puchtler,H., Waldrop,F.S., Meloan,S.N., Terry,M.S., & Conner,H.M. (1970). Methacarn (methanol-Carnoy) fixation. Practical and theoretical considerations. *Histochemie.*, 21(2), 97-116.

Pumfrey,R.A., Johnson,R.K., Cunningham,P.J., & Zimmerman,D.R. (1980). Inheritance of teat number and its relationship to maternal traits in swine. *J Anim Sci.*, 50(6), 1057-1060.

Quiniou,N., Dagorn,J., & Gaudrθ,D. (2002). Variation of piglets' birth weight and consequences on subsequent performance. *Liv Prod Sci.*, 78(1), 63-70.

R Development Core Team. R: A Language and Environment for Statistical Computing. 2005. Vienna, Austria, R foundation for Statistical Computing. Ref Type: Computer Program

Ramos,A.M., Crooijmans,R.P., Affara,N.A., Amaral,A.J., Archibald,A.L., Beever,J.E., Bendixen,C., Churcher,C., Clark,R., Dehais,P., Hansen,M.S., Hedegaard,J., Hu,Z.L., Kerstens,H.H., Law,A.S., Megens,H.J., Milan,D., Nonneman,D.J., Rohrer,G.A., Rothschild,M.F., Smith,T.P., Schnabel,R.D., Van Tassell,C.P., Taylor,J.F., Wiedmann,R.T., Schook,L.B., & Groenen,M.A. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS One.*, 4(8), e6524.

Ramos,A., Duijvesteijn,N., Knol,E., Merks,J., Bovenhuis,H., Crooijmans,R., Groenen,M., & Harlizius,B. (2011). The distal end of porcine chromosome 6p is involved in the regulation of skatole levels in boars. *BMC Genetics*, 12(1), 35.

Randall,G.C. (1978). Perinatal mortality: some problems of adaptation at birth. *Adv Vet Sci Comp Med.*, 22, 53-81.

Rathje,T.A., Rohrer,G.A., & Johnson,R.K. (1997). Evidence for quantitative trait loci affecting ovulation rate in pigs. *J Anim Sci.*, 75(6), 1486-1494.

Regnault,T.R., Galan,H.L., Parker,T.A., & Anthony,R.V. (2002). Placental development in normal and compromised pregnancies-- a review. *Placenta*, 23 Suppl A, S119-S129.

Revel,A., Helman,A., Koler,M., Shushan,A., Goldshmidt,O., Zcharia,E., Aingorn,H., & Vlodavsky,I. (2005). Heparanase improves mouse embryo implantation. *Fertil Steril.*, 83(3), 580-586.

Rice,G.E., Dantzer,V., Madsen,M.T., & Skadhauge,E. (1991). Gestational changes in electrolyte transport, electrical activity, and permeability of the porcine placenta. *J Comp Physiol B.*, 161(2), 189-198.

Roberts,R.M. & Bazer,F.W. (1988). The functions of uterine secretions. *J Reprod Fertil.*, 82(2), 875-892.

Robic,A., Dalens,M., Woloszyn,N., Milan,D., Riquet,J., & Gellin,J. (1994). Isolation of 28 new porcine microsatellites revealing polymorphism. *Mamm Genome.*, 5(9), 580-583.

Rodriguez,C., Tomas,A., Alves,E., Ramirez,O., Arque,M., Muñoz,G., Barragan,C., Varona,L., Silio,L., Amills,M., & Noguera,J.L. (2005). QTL mapping for teat number in an Iberian-by-Meishan pig intercross. *Anim Genet.*, 36(6), 490-496.

Roehe,R. & Kennedy,B.W. (1993). The influence of maternal effects on accuracy of evaluation of litter size in swine. *J Anim Sci.*, 71(9), 2353-2364.

Rohrer,G.A. (2000). Identification of quantitative trait loci affecting birth characters and accumulation of backfat and weight in a Meishan-White Composite resource population. *J Anim Sci.*, 78(10), 2547-2553.

Rohrer,G.A., Alexander,L.J., Hu,Z., Smith,T.P., Keele,J.W., & Beattie,C.W. (1996). A comprehensive map of the porcine genome. *Genome Res.*, 6(5), 371-391.

Rohrer,G.A., Alexander,L.J., Keele,J.W., Smith,T.P., & Beattie,C.W. (1994). A microsatellite linkage map of the porcine genome. *Genetics*, 136(1), 231-245.

Rohrer,G.A., Ford,J.J., Wise,T.H., Vallet,J.L., & Christenson,R.K. (1999). Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. *J Anim Sci.*, 77(6), 1385-1391.

Rosen,V. & Thies,R.S. (1992). The BMP proteins in bone formation and repair. *Trends Genet.*, 8(3), 97-102.

Rosendo,A., Druet,T., Pery,C., & Bidanel,J.P. (2010). Correlative responses for carcass and meat quality traits to selection for ovulation rate or prenatal survival in French Large White pigs. *J Anim Sci*, 88(3), 903-911.

Ross,J.W., Ashworth,M.D., White,F.J., Johnson,G.A., Ayoubi,P.J., DeSilva,U., Whitworth,K.M., Prather,R.S., & Geisert,R.D. (2007). Premature estrogen exposure alters endometrial gene expression to disrupt pregnancy in the pig. *Endocrinology*, 148(10), 4761-4773.

Rothschild,M.F. (2004). Porcine genomics delivers new tools and results: this little piggy did more than just go to market. *Genet.Res.*, 83(1), 1-6.

Rothschild,M.F. (2003). Advances in pig genomics and functional gene discovery. *Comp Funct Genomics.*, 4(2), 266-270.

Rothschild,M.F. (1996). Genetics and reproduction in the pig. *Anim Reprod Sci.*, 42, 143-151.

Rothschild,M.F., Jacobson,C., Vaske,D.A., Tuggle,C.K., Wang,L., Short,T.H., Eckardt,G.R., Sasaki,S., Vincent,A.L., McLaren,D.G., Southwood,O.I., van der Steen,H.A.M., Mileham,A.J., & Plastow,G.S. (1996). The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc Natl Acad Sci U.S.A.*, 93(1), 201-205.

Rothschild,M.F., Messer,L., Day,A., Wales,R., Short,T., Southwood,O., & Plastow,G. (2000). Investigation of the retinol-binding protein 4 (RBP4) gene as a candidate gene for increased litter size in pigs. *Mamm Genome.*, 11(1), 75-77.

Ruiz-Flores,A. & Johnson,R.K. (2001). Direct and correlated responses to two-stage selection for ovulation rate and number of fully formed pigs at birth in swine. *J Anim Sci.*, 79(9), 2286-2297.

Rutherford, K. M. D., Baxter, E. M., Ask, B., Berg, P., D'Eath, R. B., Jarvis, S., Jensen, K. K., Lawrence, A. B., Moustsen, V. A., Robson, S. K., Roehe, R., Thorup, F., Turner, S. P., and Sandøe, P. the Ethical and Welfare implications of Large Litter size in the Domestic pig. 2011. Danish centre for bioethics and Risk assessment and SAC (Scottish Agricultural College).
Ref Type: Report

Sato,S., Atsuji,K., Saito,N., Okitsu,M., Sato,S., Komatsuda,A., Mitsuhashi,T., Nirasawa,K., Hayashi,T., Sugimoto,Y., & Kobayashi,E. (2006). Identification of quantitative trait loci affecting corpora lutea and number of teats in a Meishan x Duroc F2 resource population. *J Anim Sci.*, 84(11), 2895-2901.

Seaton,G., Haley,C.S., Knott,S.A., Kearsey,M., & Visscher,P.M. (2002). QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics.*, 18(2), 339-340.

Senger,D.R. & Perruzzi,C.A. (1996). Cell migration promoted by a potent GRGDS-containing thrombin-cleavage fragment of osteopontin. *Biochim Biophys Acta.*, 1314(1-2), 13-24.

Serenius,T., Sev≤n-Aimonen,M.-L., & MΣntysaari,E.A. (2003). 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*, 81(2-3), 213-222.

Serra,J.J., Ellis,M., & Haley,C.S. (1992). Genetic components of carcass and meat quality traits in Meishan and Large White pigs and their reciprocal crosses. *Anim Sci.*, 54(01), 117-127.

Shibutani,M., Uneyama,C., Miyazaki,K., Toyoda,K., & Hirose,M. (2000). Methacarn fixation: a novel tool for analysis of gene expressions in paraffin-embedded tissue specimens. *Lab Invest.*, 80(2), 199-208.

Shimasaki,S., Zachow,R.J., Li,D., Kim,H., Iemura,S., Ueno,N., Sampath,K., Chang,R.J., & Erickson,G.F. (1999). A functional bone morphogenetic protein system in the ovary. *Proc Natl Acad Sci U.S.A.*, 96(13), 7282-7287.

Short,T.H., Rothschild,M.F., Southwood,O.I., McLaren,D.G., de,V.A., van der,S.H., Eckardt,G.R., Tuggle,C.K., Helm,J., Vaske,D.A., Mileham,A.J., & Plastow,G.S. (1997). Effect of the estrogen receptor locus on reproduction and production traits in four commercial pig lines. *J Anim Sci.*, 75(12), 3138-3142.

Sironen,A.I., Uimari,P., Serenius,T., Mote,B., Rothschild,M., & Vilkki,J. (2010). Effect of polymorphisms in candidate genes on reproduction traits in Finnish pig populations. *J Anim Sci.*, 88(3), 821-827.

Sonesson,A.K., de Greef,K.H., & Meuwissen,T.H.E. (1998). Genetic parameters and trends of meat quality, carcass composition and performance traits in two selected lines of large white pigs. *Livestock Production Science*, 57(1), 23-32.

Souza,C.J., MacDougall,C., MacDougall,C., Campbell,B.K., McNeilly,A.S., & Baird,D.T. (2001). The Booroola (FecB) phenotype is associated with a mutation in the bone morphogenetic receptor type 1 B (BMPR1B) gene. *J Endocrinol.*, 169(2), R1-R6.

Spelman,R.J. & Bovenhuis,H. (1998). Moving from QTL experimental results to the utilization of QTL in breeding programmes. *Animal Genetics*, 29(2), 77-84.

Spencer,T.E., Burghardt,R.C., Johnson,G.A., & Bazer,F.W. (2004). Conceptus signals for establishment and maintenance of pregnancy. *Anim Reprod Sci.*, 82-83, 537-550.

Spotter,A. & Distl,O. (2006). Genetic approaches to the improvement of fertility traits in the pig. *Vet J.*, 172(2), 234-247.

Spotter,A., Muller,S., Hamann,H., & Distl,O. (2009). Effect of polymorphisms in the genes for LIF and RBP4 on litter size in two German pig lines. *Reprod Domest Anim.*, 44(1), 100-105.

Srinivasan,M., Sedmak,D., & Jewell,S. (2002). Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol.*, 161(6), 1961-1971.

Stinckens,A., Mathur,P., Janssens,S., Bruggeman,V., Onagbesan,O.M., Schroyen,M., Spincemaille,G., Decuypere,E., Georges,M., & Buys,N. (2010). Indirect effect of IGF2 intron3 g.3072G>A mutation on prolificacy in sows. *Anim Genet.*, 41(5), 493-498.

Su,G., Lund,M.S., & Sorensen,D. (2007). Selection for litter size at day five to improve litter size at weaning and piglet survival rate. *J Anim Sci.*, 85(6), 1385-1392.

Sun,X., Mei,S., Tao,H., Wang,G., Su,L., Jiang,S., Deng,C., Xiong,Y., & Li,F. (2011). Microarray profiling for differential gene expression in PMSG-hCG stimulated preovulatory ovarian follicles of Chinese Taihu and Large White sows. *BMC Genomics.*, 12, 111.

Sutherland,A.E., Calarco,P.G., & Damsky,C.H. (1993). Developmental regulation of integrin expression at the time of implantation in the mouse embryo. *Development*, 119(4), 1175-1186.

Takagi,H., Shibutani,M., Kato,N., Fujita,H., Lee,K.Y., Takigami,S., Mitsumori,K., & Hirose,M. (2004). Microdissected region-specific gene expression analysis with methacarn-fixed, paraffin-embedded tissues by real-time RT-PCR. *J Histochem Cytochem.*, 52(7), 903-913.

Tayade,C., Fang,Y., & Croy,B.A. (2007). A review of gene expression in porcine endometrial lymphocytes, endothelium and trophoblast during pregnancy success and failure. *J Reprod Dev.*, 53(3), 455-463.

Thellin,O., Zorzi,W., Lakaye,B., De,B.B., Coumans,B., Hennen,G., Grisar,T., Igout,A., & Heinen,E. (1999). Housekeeping genes as internal standards: use and limits. *J Biotechnol.*, 75(2-3), 291-295.

Tomas,A., Frigo,E., Casellas,J., Ramirez,O., Ovilo,C., Noguera,J.L., & Sanchez,A. (2006). An association study between polymorphisms of the porcine bone morphogenetic protein receptor type1beta(BMPR1B) and reproductive performance of Iberian x Meishan F2 sows. *Anim Genet.*, 37(3), 297-298.

Tramontana,S., Bionaz,M., Sharma,A., Graugnard,D.E., Cutler,E.A., jmone-Marsan,P., Hurley,W.L., & Loor,J.J. (2008). Internal controls for quantitative polymerase chain reaction of swine mammary glands during pregnancy and lactation. *J Dairy Sci.*, 91(8), 3057-3066.

Tribout,T., Iannuccelli,N., Druet,T., Gilbert,H., Riquet,J., Gueblez,R., Mercat,M.J., Bidanel,J.P., Milan,D., & Le,R.P. (2008). Detection of quantitative trait loci for reproduction and production traits in Large White and French Landrace pig populations. *Genet Sel Evol.*, 40(1), 61-78.

Tuggle,C.K., Wang,Y., & Couture,O. (2007). Advances in swine transcriptomics. *Int J Biol Sci.*, 3(3), 132-152.

Tyrrell,L., Elias,J., & Longley,J. (1995). Detection of specific mRNAs in routinely processed dermatopathology specimens. *Am J Dermatopathol.*, 17(5), 476-483.

Uneyama,C., Shibutani,M., Masutomi,N., Takagi,H., & Hirose,M. (2002). Methacarn fixation for genomic DNA analysis in microdissected, paraffin-embedded tissue specimens. *J Histochem Cytochem.*, 50(9), 1237-1245.

Vallet,J.L. & Christenson,R.K. (1993). Uterine space affects placental protein secretion in swine. *Biol Reprod.*, 48(3), 575-584.

Vallet,J.L., Christenson,R.K., Trout,W.E., & Klemcke,H.G. (1998). Conceptus, progesterone, and breed effects on uterine protein secretion in swine. *J Anim Sci.*, 76(10), 2657-2670.

Vallet,J.L. & Freking,B.A. (2007). Differences in placental structure during gestation associated with large and small pig fetuses. *J Anim Sci.*, 85(12), 3267-3275.

Vallet,J.L., Freking,B.A., Leymaster,K.A., & Christenson,R.K. (2005a). Allelic variation in the secreted folate binding protein gene is associated with uterine capacity in swine. *J Anim Sci.*, 83(8), 1860-1867.

Vallet,J.L., Freking,B.A., Leymaster,K.A., & Christenson,R.K. (2005b). Allelic variation in the erythropoietin receptor gene is associated with uterine capacity and litter size in swine. *Anim Genet.*, 36(2), 97-103.

Vallet,J.L., Klemcke,H.G., & Christenson,R.K. (2002a). Interrelationships among conceptus size, uterine protein secretion, fetal erythropoiesis, and uterine capacity. *J Anim Sci.*, 80(3), 729-737.

Vallet,J.L., Leymaster,K.A., & Christenson,R.K. (2002b). The influence of uterine function on embryonic and fetal survival. *J Anim Sci.*, 80(E-Suppl_2), E115-E125.

van der Lende,T. & van Rens,B.T.T.M. (2003). Critical periods for foetal mortality in gilts identified by analysing the length distribution of mummified foetuses and frequency of non-fresh stillborn piglets. *Anim Reprod Sci.*, 75(1-2), 141-150.

van der Steen, H. A. M., Southwood, O. I., de Vries, A. G., Short, T. H., McLaren, D. G., Wei, M., and Plastow, G. S. Evidence of a new genetic marker for litter size in Meishan synthetic pigs. Proceedings of the International Conference on Animal Biotechnology , 17-20. 1997. 11-6-1997.
Ref Type: Conference Proceeding

van Rens,B.T.T.M., de Groot,P.N., & van der Lende,T. (2002). The effect of estrogen receptor genotype on litter size and placental traits at term in F2 crossbred gilts. *Theriogenology*, 57(6), 1635-1649.

van Rens,B.T.T.M., de Koning,G., Bergsma,R., & van der Lende,T. (2005). Preweaning piglet mortality in relation to placental efficiency. *J Anim Sci.*, 83(1), 144-151.

van Rens,B.T.T.M., Evans,G.J., & van der Lende,T. (2003). Components of litter size in gilts with different prolactin receptor genotypes. *Theriogenology*, 59(3-4), 915-926.

van Rens,B.T.T.M. & van der Lende,T. (2002). Litter size and piglet traits of gilts with different prolactin receptor genotypes. *Theriogenology*, 57(2), 883-893.

van Rens,B.T.T.M. & van der Lende,T. (2004). Parturition in gilts: duration of farrowing, birth intervals and placenta expulsion in relation to maternal, piglet and placental traits. *Theriogenology*, 62(1-2), 331-352.

Vandesompele,J., de Preter,K., Pattyn,F., oppe,B., van Roy,N., de Paepe,A., & peleman,F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, 3(7).

Venter,J.C., Adams,M.D., Myers,E.W., Li,P.W., Mural,R.J., Sutton,G.G., Smith,H.O., Yandell,M., Evans,C.A., Holt,R.A., Gocayne,J.D., Amanatides,P., Ballew,R.M., Huson,D.H., Wortman,J.R., Zhang,Q., Kodira,C.D., Zheng,X.H., Chen,L., Skupski,M., Subramanian,G., Thomas,P.D., Zhang,J., Gabor Miklos,G.L., Nelson,C., Broder,S., Clark,A.G., Nadeau,J., McKusick,V.A., Zinder,N., Levine,A.J., Roberts,R.J., Simon,M., Slayman,C., Hunkapiller,M., Bolanos,R., Delcher,A., Dew,I., Fasulo,D., Flanigan,M., Florea,L., Halpern,A., Hannenhalli,S., Kravitz,S., Levy,S., Mobarry,C., Reinert,K., Remington,K., bu-Threideh,J., Beasley,E., Biddick,K., Bonazzi,V., Brandon,R., Cargill,M., Chandramouliswaran,I., Charlab,R., Chaturvedi,K., Deng,Z., Francesco,V.D., Dunn,P., Eilbeck,K., Evangelista,C., Gabrielian,A.E., Gan,W., Ge,W., Gong,F., Gu,Z., Guan,P., Heiman,T.J., Higgins,M.E., Ji,R.R., Ke,Z., Ketchum,K.A., Lai,Z., Lei,Y., Li,Z., Li,J., Liang,Y., Lin,X., Lu,F., Merkulov,G.V., Milshina,N., Moore,H.M., Naik,A.K., Narayan,V.A., Neelam,B., Nusskern,D., Rusch,D.B., Salzberg,S., Shao,W., Shue,B., Sun,J., Wang,Z.Y., Wang,A., Wang,X., Wang,J., Wei,M.H., Wides,R., Xiao,C., Yan,C., Yao,A., Ye,J., Zhan,M., Zhang,W., Zhang,H., Zhao,Q., Zheng,L., Zhong,F., Zhong,W., Zhu,S.C., Zhao,S., Gilbert,D., Baumhueter,S., Spier,G., Carter,C., Cravchik,A., Woodage,T., Ali,F., An,H., Awe,A., Baldwin,D., Baden,H., Barnstead,M., Barrow,I., Beeson,K., Busam,D., Carver,A., Center,A., Cheng,M.L., Curry,L., Danaher,S., Davenport,L., Desilets,R., Dietz,S., Dodson,K., Doup,L., Ferriera,S., Garg,N., Gluecksmann,A., Hart,B., Haynes,J., Haynes,C., Heiner,C., Hladun,S., Hostin,D., Houck,J., Howland,T., Ibegwam,C., Johnson,J., Kalush,F., Kline,L., Koduru,S., Love,A., Mann,F., May,D., McCawley,S., McIntosh,T., McMullen,I., Moy,M., Moy,L., Murphy,B., Nelson,K., Pfannkoch,C., Pratts,E., Puri,V., Qureshi,H., Reardon,M., Rodriguez,R., Rogers,Y.H., Romblad,D., Ruhfel,B., Scott,R., Sitter,C., Smallwood,M., Stewart,E., Strong,R., Suh,E., Thomas,R., Tint,N.N., Tse,S., Vech,C., Wang,G., Wetter,J., Williams,S., Williams,M., Windsor,S., Winn-Deen,E., Wolfe,K., Zaveri,J., Zaveri,K., Abril,J.F., Guig+¦,R., Campbell,M.J., Sjolander,K.V., Karlak,B., Kejariwal,A., Mi,H., Lazareva,B., Hatton,T., Narechania,A., Diemer,K., Muruganujan,A., Guo,N., Sato,S., Bafna,V., Istrail,S., Lippert,R., Schwartz,R., Walenz,B., Yooseph,S., Allen,D., Basu,A., Baxendale,J., Blick,L., Caminha,M., Carnes-Stine,J., Caulk,P., Chiang,Y.H., Coyne,M., Dahlke,C., Mays,A.D., Dombroski,M., Donnelly,M., Ely,D., Esparham,S., Fosler,C., Gire,H., Glanowski,S., Glasser,K., Glodek,A., Gorokhov,M., Graham,K., Gropman,B., Harris,M., Heil,J., Henderson,S., Hoover,J., Jennings,D., Jordan,C., Jordan,J., Kasha,J., Kagan,L., Kraft,C., Levitsky,A., Lewis,M., Liu,X., Lopez,J., Ma,D., Majoros,W., McDaniel,J., Murphy,S., Newman,M., Nguyen,T., Nguyen,N., & Nodell,M. (2001). The Sequence of the Human Genome. *Science*, 291(5507), 1304-1351.

Vianna,W.L., Pinese,M.E., de Campos,R.A., Bombonato,P.P., Rodrigues,P.H., & de Sant'Anna,M.A. (2004). Relationship between prenatal survival rate at 70 days of gestation and morphometric parameters of vagina, uterus and placenta in gilts. *Reprod Domest Anim.*, 39(6), 381-384.

Vincent, A. L., Evans, G., Short, T. H., Southwood, O. I., Plastow, G. S., Tuggle, C. K., and Rothschild, M. F. The prolactin receptor gene is associated with increased litter size in pigs. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production , 15-

18. 1998.
Ref Type: Conference Proceeding

Vingborg,R.K., Gregersen,V.R., Zhan,B., Panitz,F., Hoj,A., Sorensen,K.K., Madsen,L.B., Larsen,K., Hornshoj,H., Wang,X., & Bendixen,C. (2009). A robust linkage map of the porcine autosomes based on gene-associated SNPs. *BMC Genomics.*, 10, 134.

Vlodavsky,I., Friedmann,Y., Elkin,M., Aingorn,H., Atzmon,R., Ishai-Michaeli,R., Bitan,M., Pappo,O., Peretz,T., Michal,I., Spector,L., & Pecker,I. (1999). Mammalian heparanase: gene cloning, expression and function in tumor progression and metastasis. *Nat Med.*, 5(7), 793-802.

von Wolff,M., Strowitzki,T., Becker,V., Zepf,C., Tabibzadeh,S., & Thaler,C.J. (2001). Endometrial osteopontin, a ligand of beta3-integrin, is maximally expressed around the time of the "implantation window". *Fertil Steril.*, 76(4), 775-781.

Vonnahme,K.A., Wilson,M.E., & Ford,S.P. (2002). Conceptus competition for uterine space: different strategies exhibited by the Meishan and Yorkshire pig. *J Anim Sci.*, 80(5), 1311-1316.

Vonnahme,K.A., Wilson,M.E., & Ford,S.P. (2001). Relationship between placental vascular endothelial growth factor expression and placental/endometrial vascularity in the pig. *Biol Reprod.*, 64(6), 1821-1825.

Wada,Y., Akita,T., Awata,T., Furukawa,T., Sugai,N., Inage,Y., Ishii,K., Ito,Y., Kobayashi,E., Kusumoto,H., Matsumoto,T., Mikawa,S., Miyake,M., Murase,A., Shimanuki,S., Sugiyama,T., Uchida,Y., Yanai,S., & Yasue,H. (2000). Quantitative trait loci (QTL) analysis in a Meishan x Gottingen cross population. *Anim Genet.*, 31(6), 376-384.

Walling,G.A., Archibald,A.L., Cattermole,J.A., Downing,A.C., Finlayson,H.A., Nicholson,D., Visscher,P.M., Walker,C.A., & Haley,C.S. (1998). Mapping of quantitative trait loci on porcine chromosome 4. *Anim Genet.*, 29(6), 415-424.

Wang,A., Hu,X., Li,N., & Wu,C. (2003). BMPR1B maps to chromosome 8 in swine. *Anim Genet.*, 34(4), 306.

Wang,G.K., Hu,L., Fuller,G.N., & Zhang,W. (2006). An interaction between insulin-like growth factor-binding protein 2 (IGFBP2) and integrin alpha5 is essential for IGFBP2-induced cell mobility. *J Biol Chem.*, 281(20), 14085-14091.

Wang,H. & Dey,S.K. (2006). Roadmap to embryo implantation: clues from mouse models. *Nat Rev Genet.*, 7(3), 185-199.

Waterhouse,P., Parhar,R.S., Guo,X., Lala,P.K., & Denhardt,D.T. (1992). Regulated temporal and spatial expression of the calcium-binding proteins calcyclin and OPN (osteopontin) in mouse tissues during pregnancy. *Mol Reprod Dev.*, 32(4), 315-323.

Weary,D.M., Pajor,E.A., Fraser,D., & Honkanen,A.M. (1996). Sow body movements that crush piglets: a comparison between two types of farrowing accommodation. *Applied Animal Behaviour Science*, 49(2), 149-158.

Wei,W.H., de Koning,D.J., Penman,J.C., Finlayson,H.A., Archibald,A.L., & Haley,C.S. (2007). QTL modulating ear size and erectness in pigs. *Anim Genet.*, 38(3), 222-226.

Wesolowski,S.R., Raney,N.E., & Ernst,C.W. (2004). Developmental changes in the fetal pig transcriptome. *Physiol Genomics*, 16(2), 268-274.

White,F.J., Burghardt,R.C., Hu,J., Joyce,M.M., Spencer,T.E., & Johnson,G.A. (2006). Secreted phosphoprotein 1 (osteopontin) is expressed by stromal macrophages in cyclic and pregnant endometrium of mice, but is induced by estrogen in luminal epithelium during conceptus attachment for implantation. *Reprod.*, 132(6), 919-929.

White,F.J., Ross,J.W., Joyce,M.M., Geisert,R.D., Burghardt,R.C., & Johnson,G.A. (2005). Steroid regulation of cell specific secreted phosphoprotein 1 (osteopontin) expression in the pregnant porcine uterus. *Biol Reprod.*, 73(6), 1294-1301.

Wight,T.N., Kinsella,M.G., & Qwarnstrom,E.E. (1992). The role of proteoglycans in cell adhesion, migration and proliferation. *Curr Opin Cell Biol.*, 4(5), 793-801.

Wilkie,P.J., Paszek,A.A., Beattie,C.W., Alexander,L.J., Wheeler,M.B., & Schook,L.B. (1999). A genomic scan of porcine reproductive traits reveals possible quantitative trait loci (QTLs) for number of corpora lutea. *Mamm Genome.*, 10(6), 573-578.

Williams,J.L. (2005). The use of marker-assisted selection in animal breeding and biotechnology. *Rev.Sci Tech.*, 24(1), 379-391.

Wilmut,I., Ritchie,W.A., Haley,C.S., Ashworth,C.J., & Aitken,R.P. (1992). A comparison of rate and uniformity of embryo development in Meishan and European white pigs. *J Reprod Fertil.*, 95(1), 45-56.

Wilson,M.E., Biensen,N.J., & Ford,S.P. (1999). Novel insight into the control of litter size in pigs, using placental efficiency as a selection tool. *J Anim Sci.*, 77(7), 1654-1658.

Wilson,M.E., Biensen,N.J., Youngs,C.R., & Ford,S.P. (1998). Development of Meishan and Yorkshire littermate conceptuses in either a Meishan or Yorkshire uterine environment to day 90 of gestation and to term. *Biol Reprod.*, 58(4), 905-910.

Wilson,M.E. & Ford,S.P. (2001). Comparative aspects of placental efficiency. *Reprod Suppl.*, 58, 223-232.

Wilson,M.E., Sonstegard,T.S., Smith,T.P., Fahrenkrug,S.C., & Ford,S.P. (2000). Differential gene expression during elongation in the preimplantation pig embryo. *Genesis.*, 26(1), 9-14.

Wilson,T., Wu,X.Y., Juengel,J.L., Ross,I.K., Lumsden,J.M., Lord,E.A., Dodds,K.G., Walling,G.A., McEwan,J.C., O'Connell,A.R., McNatty,K.P., & Montgomery,G.W. (2001). Highly Prolific Booroola Sheep Have a Mutation in the Intracellular Kinase Domain of Bone Morphogenetic Protein IB Receptor (ALK-6) That Is Expressed in Both Oocytes and Granulosa Cells. *Biol Reprod.*, 64(4), 1225-1235.

Wise,T., Klindt,J., Howard,H.J., Conley,A.J., & Ford,J.J. (2001). Endocrine relationships of Meishan and White composite females after weaning and during the luteal phase of the estrous cycle. *Journal of Animal Science*, 79(1), 176-187.

Wise,T., Roberts,A.J., & Christenson,R.K. (1997). Relationships of light and heavy fetuses to uterine position, placental weight, gestational age, and fetal cholesterol concentrations. *Journal of Animal Science*, 75(8), 2197-2207.

Wu,M.C., Hentzel,M.D., & Dziuk,P.J. (1988). Effect of stage of gestation, litter size and uterine space on the incidence of mummified fetuses in pigs. *J Anim Sci.*, 66(12), 3202-3207.

Yang,G., Ren,J., Li,S., Mao,H., Guo,Y., Zou,Z., Ren,D., Ma,J., & Huang,L. (2008). Genome-wide identification of QTL for age at puberty in gilts using a large intercross F2 population between White Duroc x Erhualian. *Genet Sel Evol.*, 40(5), 529-539.

Yasue,H., Mikawa,S., Uenishi,H., & Wada,Y. (1999). Analysis of allele segregation distortion in a swine resource family. *Anim Biotechnol.*, 10(3), 147-152.

Youngs,C.R., Christenson,L.K., & Ford,S.P. (1994). Investigations into the control of litter size in swine: III. A reciprocal embryo transfer study of early conceptus development. *J Anim Sci.*, 72(3), 725-731.

Zhang,J., Xiong,Y., Zuo,B., Lei,M., Jiang,S., Li,F., Zheng,R., Li,J., & Xu,D. (2007). Detection of quantitative trait loci associated with several internal organ traits and teat number trait in a pig population. *J Genet Genomics.*, 34(4), 307-314.

Zhang,Q., Wrana,J.L., & Sodek,J. (1992). Characterization of the promoter region of the porcine opn (osteopontin, secreted phosphoprotein 1) gene. Identification of positive and negative regulatory elements and a 'silent' second promoter. *Eur J Biochem.*, 207(2), 649-659.

Zhou,Q.Y., Fang,M.D., Huang,T.H., Li,C.C., Yu,M., & Zhao,S.H. (2009). Detection of differentially expressed genes between Erhualian and Large White placentas on day 75 and 90 of gestation. *BMC Genomics.*, 10, 337.

Appendix 1

List of primers of microsatellites genotyped previously in Archibald's Laboratory across the whole genome (except chromosome 8) and used in this study for the linkage maps. The primers genotyped in the present study are also excluded from this list. The chromosome (SSC), the name of the marker, the label used, the forward and reverse sequence of the primers, the product size, the annealing temperature (Tm) and the number of animals with genotypes results are indicated in the table.

SSC	Marker name	Primers	Primer sequence 5' - 3'	product size bp	T°m	Genotype counts
1	**CGA**	CGA-FAM	GAACTTTCACATCCCTAAGGTCGT	277-289	55	303
		CGA-R	ATAGACATTATGTCCGTTGCTGAT			
	S0008	S0008-VIC	GAGGCAGTGTGTTCTATTCA	186-190	56	305
		S0008-R	GCCATGTGTAAAGTGTTGCT			
	S0082	S0082-FAM	CAGAAAATAAACTTGTCTAACTTG	154-182	55	302
		S0082-R	AACCCTGTTTCATATCATTAAGCC			
	S0122	S0122-FAM	GCCCACAGTGGGAACTTC	170-196	60	301
		S0122-R	AATATACAAGCTTGGGGTCACA			
	S0155	S0155-FAM	TGTTCTCTGTTTCTCCTCTGTTTG	148-166	55	296
		S0155-R	AAAGTGGAAAGAGTCAATGGCTAT			
	SW1301	SW1301-NED	TGGATAAGCAATGAGGTCCC	144-176	58	304
		SW1301-R	TAGTGGATTTATAATGTGCTAACCC			
	SW1515	SW1515-PET	CTCCGGTTTCCATTTGTGG	115-148	60	306
		SW1515-R	GATCCCTGCCCCCAACAC			
	SW64	SW64-NED	AGACCAAGGGCCATGAGAG	130-160	56	301
		SW64-R	TTCCACGTGATGTGGGATAG			
2	**S0036**	S0036-FAM	AGTGACGTGAGGGTCTGCTCCTC	114-128	65	306
		S0036-R	ATGGACGGTGGATTCACAGCC			

2	**S0091**	S0091-NED	TCTACTCCAGGAGATAAGCCAGAT	148-170	55	306
		S0091-R	CAGTGACTCCATGCACAGTTATGA			
	S0226	S0226-FAM	GCACTTTTAACTTTCATGATACTCC	181-206	55	306
		S0226-R	GGTTAAACTTTTNCCCCAATACA			
	S0378	S0378-NED	TGGGACCTAATTGTTGAGACAG	96-114	55	304
		S0378-R	ACTGAGCCACAATGAAGAGAAC			
	SW1026	SW1026-NED	TGGAGAGGCAATGCTGTATG	97-123	60	306
		SW1026-R	GTATTTCACCTGCAGCTCCC			
	SW1695	SW1695-FAM	ATAAGGGAAATCAGGCTGAGC	169-207	65	306
		SW1695-R	TCCCAGGAGCTACCATATGC			
	SW1879	SW1879-NED	AGACACATGCACATGTGTTTTAC	182-196	56	306
		SW1879-R	AGCATTTGTTTCTGGTTACTTTTAG			
	SW240	SW240-NED	AGAAATTAGTGCCTCAAATTGG	94-120	58	306
		SW240-R	AAACCATTAAGTCCCTAGCAAA			
	SW2443	SW2443-VIC	GAGCACAGAAGATTTTTAGGGC	200-214	62	304
		SW2443-R	TTAGTTTTCTCCTGGGCTGTG			
	SW256	SW256-NED	ACAAAAGCTTTTGGAGAACTCG	91-120	62	293
		SW256-R	TAGCATAGGAACAGGTGCAGC			
	SW395	SW395-NED	TTCCAAGGTTATGGAGATATCC	139-163	55	306
		SW395-R	GATCCCTACCTCACACCACA			
3	**S0002**	S0002-PET	GAAGCCAAAGAGACAACTGC	189-209	62	306
		S0002-R	GTTCTTTACCCACTGAGCCA			
	S0167	S0167-PET	AAACTCCAATTTCCATAACATAGG	200-226	55	306
		S0167-R	CTTCATATATGTGCTAAGACTTCT			
	SW2527	SW2527-PET	CTACCCCCAAGTATAAATTGATTTC	140-153	58	306
		SW2527-R	CACACTTGCACTACTATTGTTAAGC			

3	**SW274**	SW274-FAM	CGCACAGCGACATCTTTTTA	107-145	60	305
		SW274-R	AAGTGCAGCCCTAAAAAGACA			
	SW590	SW590-NED	GCATTGACCAGGGTCAGG	182-276	60	306
		SW590-R	ATTTGCTGAGAGATAAGGTGCC			
	SW72	SW72-FAM	ATCAGAACAGTGCGCCGT	101-115	58	306
		SW72-R	TTTGAAAATGGGGTGTTTCC			
	SW902	SW902-NED	ATCAGTTGGAAATGATGGCC	195-208	60	306
		SW902-R	CTTGCCTCAAAGAGTTGTAAGG			
4	**S0001**	S0001-FAM	TGGATGGGTCTCATTCTCAG	178-190	55	306
		S0001-R	TGATTCCTAGCCTGAGAAGC			
	S0023	S0023-NED	TCTTTGAACTAAAATATAGACTC	80-105	50	306
		S0023-R	TTCTCCAAACTCTGTCACAC			
	S0073	S0073-FAM	ACTGAAACAGGAATTCAGATCC	105-123	59	306
		S0073-R	TGAAGTATTATGGCATCATGGA			
	S0097	S0097-NED	GACCTATCTAATGTCATTATAGT	207-244	55	305
		S0097-R	TTCCTCCTAGAGTTGACAAACTT			
	S0214	S0214-NED	CCCTGCAAGCGTTCATCTCA	121-138	55	306
		S0214-R	GGCTGTGCCAAGTCCATTAG			
	S0217	S0217-FAM	GATCCCGCATTACTGTAGCTG	208-280	58	306
		S0217-R	GCCTCCTCATCTGGGGTC			
	S0227	S0227-VIC	GATCCATTTATAATTTTAGCACAAAGT	230-257	55	306
		S0227-R	GCATGGTGTGATGCTATGTCAAGC			
	S0301	S0301-FAM	CCGTCTTACTTAGGATGTTT	252-263	55	305
		S0301-R	TGATGTGTTTATGTGTTTGA			
	SW445	SW445-FAM	CCTCCCTGGCACTCATTG	181-208	58	305
		SW445-R	CACACACACAAGCAGGTGC			

5	**DAGK**	DAGK-NED	CTATTCCCCATGAACCCATG	120-131	55	306
		DAGK-R	TCCCAGTGGGAAAAAAAAGT			
	GDF11	GDF11- NED	CCACTTCTTGTGACTATGTGCAAG	113-136	58	305
		GDF11-R	GCCATCAGACCATATGGCC			
	IGF1	IGF1-FAM	GCTTGGATGGACCATGTTG	197-206	58	303
		IGF1-R	CATATTTTTCTGCATAACTTGAACCT			
	S0005	S0005-NED	TCCTTCCCTCCTGGTAACTA	202-246	58	306
		S0005-R	GCACTTCCTGATTCTGGGTA			
	SW1954	SW1954-NED	GATCGAACCCACACCACAG	168-197	58	301
		SW1954-R	TCATTTGGAATAAAGGGATTTC			
	SW413	SW413-NED	CAGACACACACCCCAGTGTC	158-187	62	306
		SW413-R	AGGTCCAACCCTCCTGATG			
	SW967	SW967-VIC	AGCAGACTGTTCATCTGTTCAG	0-96-116	58	305
		SW967-R	GGGGCAGCTGAAAAGTCC			
	SWR453	SWR453-FAM	TCTGGACTTGCTGTGACTGTG	168-200	58	305
		SWR453-R	TTGAATTTTTTTCATGGAAACC			
6	**S0031**	S0031-PET	AATGTCCACATGGTTTTATG	113-130	50	305
		S0031-R	GGGCTTAGATAACTTATGTT			
	S0035	S0035-VIC	GGCCGTCTTATACTCTCAGCATA	176-186	65	306
		S0035-R	CCAAATAAACAGCAGGCAGCCT			
	S0220	S0220-FAM	CCAGCAAGTTGAGGAGCCCAGA	143-158	55	306
		S0220-R	AGAGGATCGAAGGAACAAGAGGAA			
	S0228	S0228-NED	GGCATAGGCTGGCAGCAACA	221-244	55	305
		S0228-R	AGCCCACCTCATCTTATCTACACT			
	SW1057	SW1057-FAM	TCCCCTGTTGTACAGATTGATG	136-200	58	306
		SW1057-R	TCCAATTCCAAGTTCACTAGC			

6	**SW122**	SW122-FAM	CAAAAAAGGCAAAAGATTGACA	110-133	58	306
		SW122-	TTGTCTTTTTATTTTGCTTTTGG			
	SW316	SW316-PET	TTCTCCAGCCATCATGAGTG	133-162	58	306
		SW316-R	AATGACCATTCCTGAGGCTG			
	SW71	SW71-PET	GATCACCCTTATCCCCATTC	87-113	55	306
		SW71-R	TAGAAACACCATCATCCATTCA			
7	**BMP5**	BMP5-FAM	TTTCGAAAGAGACTAAAAC	187-212	58-35	306
		BMP5-R	AGGCACAGAGAAGGACTGGA			
	DAXX	DAXX-NED	GTGTCAGCAGGCAGGAAGA	182-196	58	306
		DAXX-R	GTGGCATAGGTTGGTGGC			
	S0025	S0025-FAM	TCTCCCTTCCCTCCATCTCT	104-120	55	306
		S0025-R	CTCCATCAGCCAAAAACATT			
	S0066	S0066-NED	ACATTTAAGGTGAAGCAGCAAGTG	134-198	63	306
		S0066-R	TGTCATCAACATTGAGAATTGGTG			
	S0101	S0101-FAM	GAATGCAAAGAGTTCAGTGTAGG	195-224	59	306
		S0101-R	GTCTCCCTCACACTTACCGCAG			
	S0102	S0102-FAM	GTCAAGCAAACTTCCACGTCT	114-146	55	306
		S0102-R	ATTTTGTGCCAAATGCATTGTG			
	SW1856	SW1856-NED	TCATTCCAAACACACAGAGTCC	176-210	58	306
		SW1856-R	TTGTATGGTATCCTGTGATGCC			
	SW2019	SW2019-PET	ATGATGCGAACCTGGAACTC	126-143	58	306
		SW2019-R	TATGTGTAACTTGGTCCCATGC			
	SW2155	SW2155-VIC	AGGGTGACAGACCAGAATGG	135-151	62	305
		SW2155-R	TCTGGGTCACAGGGAATTTC			
	SW632	SW632-NED	TGGGTTGAAAGATTTCCCAA	151-180	58	306
		SW632-R	GGAGTCAGTACTTTGGCTTGA			

7	**SW764**	SW764-FAM	TAGCAGATTGTTTAGCCTCTGTG	112-128	60	306
		SW764-R	AAGCATCTTTTCTAAGCACAACA			
	TNFB	TNFB-PET	CTGGTCAGCCACCAAGATTT	174-213	55	306
		TNFB-R	GGAAATGAGAAGTGTGGAGACC			
8	**SW2410**	SW2410-F	ATTTGCCCCCAAGGTATTTC	108 124	55	335
		SW2410-R	CAGGGTGTGGAGGGTAGAAG			
	HD-1	HD-1-F	GAGCGTGGACAGAGTGAACG	500	63	333
		HD-1-R	ACAGACACACGCTCCATGGC			
	SW2611	SW2611-F	CTTGTTTCCCGCAGTCTCTC	141-178	58	306
		SW2611-R	GTGTGTTCCAGATGAACCTGG			
	SW905	SW905-F	ATCCCAACCTTCTTTCAAAGG	151-125	60	310
		SW905-R	TCCAGTGGCAGAACAACATG			
	QDPR-1	QDPR-1-F	TGGGTGAAGGTGGATGGTTTTC	220	61	334
		QDPR-1-R	AGCTGAAGTCCGCGTGGG			
	SLIT2	SLIT2-F	ACTACCAAAAGCAGCAGGGCTATG	300	61	334
		SLIT2-R	CAAATATTTCCACTATGAAGCATTCA			
	SW268	SW268-F	CTGATTCACTTTCATTCGAGAA	118-168	60	90
		SW268-R	AGCCCTTCCCTTAATATAACCC			
	SW7	SW7-F	TAACCATGCTTTTCCTAGGTGG	112-89	65	329
		SW7-R	CCAGAGCTGAGTAAAAAGGTCA			
	SULTE1	SULTE1-F	CCCAGCCTCAGCAATAGTATTAATA			330
		SULTE1-R	ACTTATGTCTTCGTATCTAG			
	S0017	S0017-F	CTAGGAGAAAATCTGAGGTT	154 175		306
		S0017-R	GTTTGAATGGAGGTGCTGTA			

8	**AREG**	AREG-F	CCAAAAGAAAGAAAAAGGGAGGCA			335
		AREG-R	GCGGCTTTTCCCCACATCGTTCACC			
	FGG-2	FGG-2-F	GTTTGTAGCATGTTAAAAATTTCGC			333
		FGG-2-R	ATTTCCAGACCCATCAATTTCA			
	S0225	S0225-F	GCTAATGCCAGAGAAATGCAGA			331
		S0225-R	CAGGTGGAAAGAATGGAATGAA			
	SW61	SW61-F	GAGAGGGATGAGCACTCTGG	262-238	62	327
		SW61-R	AGAGCATTCCAGGCTTCTA			
	SPP1-1	SPP1-1-F	TTAGGGGACCCAGAGATG	209		325
		SPP1-1-R	AGATGTGTCATGAGGTTTGTGC			
	SPP1-4	SPP1-4-F	TCACCGATTTCCCCACCGACAC		65	330
		SPP1-4-R	TGGCTGCGGGTTTCCACACTG			
	SPP1-5	SPP1-5-F	GATGGGAGCAATGAGCATTC	160	57	330
		SPP1-5-R	AAGACGCACTCTCTAATTCATGAGA			
	IBSP	IBSP-F	GAGTACAGCTACTACAAAGGGCGCA	382	61	331
		IBSP-R	GGGTGCAAACACCTTAAAAATACCG			
	S0178	S0178-F	TAGCCTGGGAACCTCCAGACGCTG			334
		S0178-R	GGCACCACCAATCTGCAATCCAGT			
9	**APOA1**	APOA1-VIC	CAGCTCCTCGGTCTATCTGG	127-148	55	306
		APOA1-R	GGTTGCTGGGCAGACCTCAGCCTA			
	S0295	S0295-PET	TACTGCTGAGGCAAAGGA	228-260	52	304
		S0295-R	GCCTAAAAAGACCAAAGAA			
	SW1677	SW1677-FAM	TTCTGTTTTGGCTCTAGAGGAG	97-135	58	306
		SW1677-R	CAGATACACCCATGTTCCAGC			

9	**SW174**	SW174-NED	TGCTCTAATCTACCCGGGTC	122-131	58	305
		SW174-R	TCATGCTATTTTGTTCCAGATG			
	SW749	SW749-FAM	TTCCCAAACCAACCAAAGAG	101-113	58	306
		SW749-R	AGGAACTTGCCAAAATCACG			
	SW911	SW911-FAM	CTCAGTTCTTTGGGACTGAACC	149-173	60	306
		SW911-R	CATCTGTGGAAAAAAAAAGCC			
	SW983	SW983-NED	GCAGTCCCACTCTTAGGTATATATCC	88-129	60	306
		SW983-R	ATAATGCTGCTATGAACACTGTAGTG			
10	**S0070**	S0070-NED	GGCGAGCATTTCATTCACAG	261-299	62	306
		S0070-R	GAGCAAACAGCATCGTGAGC			
	SW1041	SW1041-PET	GGAGAATTCCCAAAGTTAATAGG	95-106	58	306
		SW1041-R	ATCAGAAAATGGTCAACAGTTCA			
	SW443	SW443-FAM	ACAAAGGCCAAGCCACATAC	105-151	60	304
		SW443-R	TCACCAGGTTTCTGGGTTTC			
	SW830	SW830-FAM	AAGTACCATGGAGAGGGAAATG	179-205	62	306
		SW830-R	TGAGTGCAACCGTGGTTAGG			
10	**SW951**	SW951-PET	TTTCACAACTCTGGCACCAG	122-136	58	306
		SW951-R	GATCGTGCCCAAATGGAC			
	SWR67	SWR67-NED	GTCCTCATGGAGACTAGTTGGG	126-154	60	305
		SWR67-R	TCC ATG CCA TGG ACA CAG			
11	**S0230**	S0230-FAM	AACAGCCCAAGTGCCCATT	296-328	55	303
		S0230-R	TCCCCCTCCACTTCCTTTC			
	S0385	S0385-VIC	AGTTCAGAAGCTGTTGCT	147-200	55	305
		S0385-R	CTATTAGGCTGGAGGGTTG			
	SW151	SW151-VIC	TTCCCCTATGATGAGATGGC	195-208	60	306
		SW151-R	GGTGTGGCCCTCAAAAGG			

11	**SW1632**	SW1632-VIC	GTTTGACAGATAAGGCTCCTGC	185-225	58	305
		SW1632-R	ACACGCTCCCTAATCCCC			
	SW703	SW703-NED	AAGATGAAGCAGGAACTCAAGG	129-143	58	304
		SW703-R	CTTGATGGCTTTACTGTTCACC			
12	**S0090**	S0090-FAM	CCAAGACTGCCTTGTAGGTGAATA	241-260	55	300
		S0090-R	GCTATCAAGTATTGTACCATTAGG			
	S0143	S0143-NED	ACTCACAGCTTGTCCTGGGTGT	142-172	50	304
		S0143-R	CAGTCAGCAGGCTGACAAAAAC			
	SW2490	SW2490-FAM	TGTTTGTCTGTCTGTCTCTCTTCC	119-176	58	306
		SW2490-R	TGTGCTTTTCAGAGGCAGG			
	SW874	SW874-FAM	AAAAGAACCCAACTACAGCAGC	190-220	60	306
		SW874-R	TTTATGAGGGTATCCTGACACC			
	SW957	SW957-PET	AGGAAGTGAGCTCAGAAAGTGC	113-157	58	306
		SW957-R	ATGGACAAGCTTGGTTTTCC			
13	**S0068**	S0068-NED	CCTTCAACCTTTGAGCAAGAAC	206-263	62	302
		S0068-R	AGTGGTCTCTCTCCCTCTTGCT			
	S0076	S0076-NED	TGAGATCTGCTGTATCCATC	145-184	55	302
		S0076-R	CTCTATTAGTTTGGTTGGCC			
	S0215	S0215-PET	TAGGCTCAGACCCTGCTGCAT	135-196	55	306
		S0215-R	TGGGAGGCTGAAGGATTGGGT			
	SW1056	SW1056-FAM	TTTCTGGTGTACAGCAAAGTGA	149-184	62	305
		SW1056-R	GGTGGTTGGTCCTCAAAAACA			
	SW1378	SW1378-NED	ACCACACGTCTAATTGAAGAGC	91-104	60	304
		SW1378-R	TAAATCACAACTTTTGGGGATG			

13	**SW398**	SW398-PET	AAGTGCCAATGCTTTGTTCC	166-189	55	305
		SW398-R	CGGAGGAGAAATAAGGGTAGC			
	SW769	SW769-PET	TCTGCTATGTGGGAAGAATGC	106-139	60	306
		SW769-R	GGTATGACCAAAAGTCCTGGG			
14	**S0007**	S0007-PET	TTACTTCTTTGGATCATGTC	160-196	56	306
		S0007-R	GTCCCTCCTCATAATTTCTG			
	SW1557	SW1557-FAM	TGCTCTAATCTACCCGGGTC	76-110	58	305
		SW1557-R	CCACCCCACTCCCTTCTG			
	SW210	SW210-FAM	TCATCACCATCATACCAAGATG	220-254	60	304
		SW210-R	AATTCTGCCAAGAAGAGAGCC			
	SW2496	SW2496-VIC	TATAGCATTTGGATGTTCCACG	183-267	55	306
		SW2496-R	GCCCAAATAAAGTGGTCTATGC			
	SW2515	SW2515-NED	CCATCTCATCCAGAAACATCC	89-105	60	306
		SW2515-R	AGGATGCTGAGGTGTTAGGC			
	SW295	SW295-FAM	ACCTGCCAGAGTTGTGGC	116-134	62	306
		SW295-R	AAGAGTTTCATTTCTCCCATCC			
	SW761	SW761-NED	CTTTGCTCCCCATTAAGCTG	136-170	58	306
		SW761-R	TCTAGCAAATGTCTGAGATGCC			
	SW857	SW857-FAM	TGAGAGGTCAGTTACAGAAGACC	139-164	58	306
		SW857-R	GATCCTCCTCCAAATCCCAT			
	SWC27	SWC27-VIC	CTGAGACTGTGCTGCTCACTG	126-165	58	306
		SWC27-R	CCATTTTCCAAAAACATGGG			
15	**S0148**	S0148-VIC	TGGTGTAGGCCTGCAGTTGA	140-186	55	306
		S0148-R	CCATCCATTGCTACTGGCAC			

15	**S0149**	S0149-PET	ATTGGCTCATGAACCACCATC	256-314	55	305
		S0149-R	GAGTTACTAATTGCCTCAGAG			
	S0355	S0355-FAM	TCTGGCTCCTACACTCCTTCTTGATG	236-282	55	306
		S0355-R	TTGGGTGGGTGCTGAAAAATAGGA			
	SW1119	SW1119-PET	CAACCTCAAAAATGGAGAAAGG	144-179	60	306
		SW1119-R	GTTCTTGCGGTGTTTGGC			
	SW936	SW936-FAM	TCTGGAGCTAGCATAAGTGCC	79-113	58	306
		SW936-R	GTGCAAGTACACATGCAGGG			
	SW964	SW964-NED	GTGGTTCCTCTATGCAGAGTCC	217-248	60	306
		SW964-R	ATGTGATGAAACATGATGGAGG			
16	**S0026**	S0026-PET	AACCTTCCCTTCCCAATCAC	80-118	58	306
		S0026-R	CACAGACTGCTTTTTACTCC			
	SW1897	SW1897-VIC	GTGCCGTGGCAGGAACTC	156-165	62	303
		SW1897-R	ACTGCCATTTGTTTTCAAAGTG			
	SW403	SW403-PET	GTGTATGTTCATGCATGGGTG	104-117	55	306
		SW403-R	GTCTCTGCTTTGCTTGCATG			
	SW742	SW742-FAM	AATTCTACTTCTGGGGAGAGGG	193-224	60	304
		SW742-R	CTTTTGGGAACATTTCTGCC			
17	**S0296-2**	S0296-2-PET	TGAAAAATAACAAGAACCAC	161-181	50	306
		S0296-2-R	AAAAGCAAATAATGATAATAG			
	S0359	S0359-NED	GGAGCTTCCATGTNCTGCAGGT	232-292	58	306
		S0359-R	GCAGAGCCTTNNGAGATTTCAGAT			
	SW2431	SW2431-PET	CTTCCCAGGATGTTGTCTAAAC	142-174	60	305
		SW2431-R	CATGGTGCACACTTAGTGGG			
	SW335	SW335-NED	GAGTATGGGGAAAGCCACG	98-113	53	80
		SW335-R	CCATCAACAAACTGTATGCACC			

18	**S0062**	S0062-PET	AAGATCATTTAGTCAAGGTCACAG	144-198	58	306
		S0062-R	TCTGATAGGGAACATAGGATAAAT			
	S0120	S0120-VIC	GCCTAAGTAGAATTAAGCACAAGG	154-174	55	305
		S0120-R	GTGCTCTCACTGCCTTCATATACC			
	S0306	S0306-PET	TGGAGACTCCAGTGGGAATG	190-236	55	301
		S0306-R	TTCTTGACAGCGAGTTGGC			
	SW1682	SW1682-VIC	GCCAATAAAACTATTCATTTGGG	142-202	58	306
		SW1682-R	AACGAGCTAGACTTTAAAGGGTG			
	SW1808	SW1808-VIC	CCAAAAAAGTGGACTGTAAGCC	104-154	60	305
		SW1808-R	TACGGATGGATGGAGACAGG			
	SW1984	SW1984-VIC	TTTTTAGTGTCCAAGGAGGTCC	100-159	55	306
		SW1984-R	GGAGCACTAATAGACCACCACC			
	SW2540	SW2540-FAM	AAATTAGGTTCTCCACGGAGC	80-108	55	306
		SW2540-R	AAACACCTAACCAGGTCACACC			
	SW787	SW787-FAM	CTGGAGCAGGAGAAAGTAAGTTC	150-166	60	305
		SW787-R	GGACAGTTACAGACAGAAGG			
X	**S0218**	S0218-FAM	GTGTAGGCTGGCGGTTGT	166-184	58	306
		S0218-R	CCCTGAAACCTAAAGCAAAG			
	SW1943	SW1943-VIC	ATTCCCCTTGACACATTAATGG	94-118	62	305
		SW1943-R	TATGGCTGAGTAGTATTCCATTTTG			
	SW2456	SW2456-NED	GAGCAACCTTGAGCTGGAAC	180-216	60	305
		SW2456-R	AATGTGATTGATGCTGTGAAG			

Appendix 2

Positional candidate genes for SSC8 QTL with effects on female reproductive traits.

This list of positional candidate genes was extracted from Ensembl release 61. This list of protein-coding genes located between 90 – 120 Mbp on SSC8 was extracted using the Ensembl Biomart data-mining tools. The positions of genetic markers used in the QTL analyses are shown in red and the genes with a known function in reproduction are highlighted.

Gene Start (bp)	HGNC symbol	Description
89,920,138		
90,035,478	FABP2	fatty acid binding protein 2, intestinal [Source:HGNC Symbol;Acc:3556]
90,060,686		cGMP-dependent 3',5'-cyclic phosphodiesterase 5A Fragment [Source:UniProtKB/TrEMBL;Acc:A8DRG4]
90,090,813		
90,168,377		
90,208,385	SYNPO2	synaptopodin 2 [Source:HGNC Symbol;Acc:17732]
90,507,160	SEC24D	SEC24 family, member D (S. Cerevisiae) [Source:HGNC Symbol;Acc:10706]
90,573,759	METTL14	methyltransferase like 14 [Source:HGNC Symbol;Acc:29330]
90,700,850	PRSS12	protease, serine, 12 (neurotrypsin, motopsin) [Source:HGNC Symbol;Acc:9477]
90,802,974	NDST3	N-deacetylase/N-sulfotransferase (309hospha glucosaminyl) 3 [Source:HGNC Symbol;Acc:7682]
90,898,795	RAB40C	RAB40C, member RAS oncogene family [Source:HGNC Symbol;Acc:18285]
91,852,366		
92,198,460		
92,246,263	LARP7	La ribonucleoprotein domain family, member 7 [Source:HGNC Symbol;Acc:24912]
92,272,536	C4orf21	chromosome 4 open reading frame 21 [Source:HGNC Symbol;Acc:25654]
93,037,208	UGT8	UDP glycosyltransferase 8 [Source:HGNC Symbol;Acc:12555]
93,408,832		
93,598,105	CAMK2D	calcium/calmodulin-dependent protein kinase II delta [Source:HGNC Symbol;Acc:1462]

93,845,914		
93,935,642	ANK2	ankyrin 2, neuronal [Source:HGNC Symbol;Acc:493]
94,159,012	S0225	
94,506,457		
94,658,948	ALPK1	alpha-kinase 1 [Source:HGNC Symbol;Acc:20917]
94,766,037	AP1AR	adaptor-related protein complex 1 associated regulatory protein [Source:HGNC Symbol;Acc:28808]
94,802,669	TIFA	TRAF-interacting protein with forkhead-associated domain [Source:HGNC Symbol;Acc:19075]
94,868,726	C4orf32	chromosome 4 open reading frame 32 [Source:HGNC Symbol;Acc:26813]
95,462,447	S0794	
95,915,280	PITX2	paired-like homeodomain 2 [Source:HGNC Symbol;Acc:9005]
96,061,729	ENPEP	glutamyl aminopeptidase (aminopeptidase A) [Source:HGNC Symbol;Acc:3355]
96,344,893		
96,469,555		
96,522,141	LRIT3	leucine-rich repeat, immunoglobulin-like and transmembrane domains 3 [Source:HGNC Symbol;Acc:24783]
96,544,378	RRH	retinal pigment epithelium-derived rhodopsin homolog [Source:HGNC Symbol;Acc:10450]
96,564,650	GAR1	GAR1 ribonucleoprotein homolog (yeast) [Source:HGNC Symbol;Acc:14264]
96,596,051	CFI	complement factor I [Source:HGNC Symbol;Acc:5394]
96,645,687	PLA2G12A	phospholipase A2, group XIIA [Source:HGNC Symbol;Acc:18554]
96,724,720	CASP6	caspase 6, apoptosis-related cysteine peptidase [Source:HGNC Symbol;Acc:1507]
96,737,112	CCDC109B	coiled-coil domain containing 109B [Source:HGNC Symbol;Acc:26076]
96,788,595	SEC24B	SEC24 family, member B (S. Cerevisiae) [Source:HGNC Symbol;Acc:10704]
97,178,019	COL25A1	collagen, type XXV, alpha 1 [Source:HGNC Symbol;Acc:18603]
97,368,295	AGXT2L1	alanine-glyoxylate aminotransferase 2-like 1 [Source:HGNC Symbol;Acc:14404]
97,440,140	OSTC	oligosaccharyltransferase complex subunit [Source:HGNC Symbol;Acc:24448]
97,472,538	RPL34	ribosomal protein L34 [Source:HGNC Symbol;Acc:10340]

97,521,499		
97,570,765	KS192	
97,663,451	LEF1	lymphoid enhancer-binding factor 1 [Source:HGNC Symbol;Acc:6551]
97,663,638		
97,786,277	HADH	hydroxyacyl-CoA dehydrogenase [Source:HGNC Symbol;Acc:4799]
97,852,755	CYP2U1	cytochrome P450, family 2, subfamily U, polypeptide 1 [Source:HGNC Symbol;Acc:20582]
97,892,312	SGMS2	sphingomyelin synthase 2 [Source:HGNC Symbol;Acc:28395]
98,741,666	DKK2	dickkopf homolog 2 (Xenopus laevis) [Source:HGNC Symbol;Acc:2892]
98,804,714	SW763	
99,228,451	TBCK	TBC1 domain containing kinase [Source:HGNC Symbol;Acc:28261]
99,545,592	GSTCD	glutathione S-transferase, C-terminal domain containing [Source:HGNC Symbol;Acc:25806]
99,647,651	ARHGEF38	Rho guanine nucleotide exchange factor (GEF) 38 [Source:HGNC Symbol;Acc:25968]
99,965,638	TET2	tet oncogene family member 2 [Source:HGNC Symbol;Acc:25941]
100,101,794		
100,648,256	CXXC4	CXXC finger protein 4 [Source:HGNC Symbol;Acc:24593]
100,864,091	TACR3	tachykinin receptor 3 [Source:HGNC Symbol;Acc:11528]
101,017,533	NHEDC2	Na+/H+ exchanger domain containing 2 [Source:HGNC Symbol;Acc:25143]
101,056,811		
101,151,372	BDH2	3-hydroxybutyrate dehydrogenase, type 2 [Source:HGNC Symbol;Acc:32389]
101,178,233	CISD2	CDGSH iron 311hospha domain 2 [Source:HGNC Symbol;Acc:24212]
101,229,430		Ubiquitin conjugating enzyme Fragment [Source:UniProtKB/TrEMBL;Acc:Q29170]
101,277,552	MANBA	mannosidase, beta A, lysosomal [Source:HGNC Symbol;Acc:6831]
101,415,843		
101,537,831	NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 [Source:HGNC Symbol;Acc:7794]
101,810,629	SLC39A8	solute carrier family 39 (zinc transporter), member 8 [Source:HGNC Symbol;Acc:20862]
102,019,884		

102,084,531	BANK1	B-cell scaffold protein with ankyrin repeats 1 [Source:HGNC Symbol;Acc:18233]
102,532,923	PPP3CA	protein 312hosphatise 3, catalytic subunit, alpha isozyme [Source:HGNC Symbol;Acc:9314]
103,062,221	EMCN	endomucin [Source:HGNC Symbol;Acc:16041]
103,112,672	DDIT4L	DNA-damage-inducible transcript 4-like [Source:HGNC Symbol;Acc:30555]
103,264,180	DAPP1	dual adaptor of phosphotyrosine and 3-phosphoinositides [Source:HGNC Symbol;Acc:16500]
103,293,154	MAPKSP1	MAPK scaffold protein 1 [Source:HGNC Symbol;Acc:15606]
103,316,645	DNAJB14	DnaJ (Hsp40) homolog, subfamily B, member 14 [Source:HGNC Symbol;Acc:25881]
103,369,419	H2AFZ	H2A histone family, member Z [Source:HGNC Symbol;Acc:4741]
103,485,922	MTTP	microsomal triglyceride transfer protein [Source:HGNC Symbol;Acc:7467]
103,551,051		
103,583,916	C4orf17	chromosome 4 open reading frame 17 [Source:HGNC Symbol;Acc:25274]
103,678,942	ADH7	alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide [Source:HGNC Symbol;Acc:256]
103,718,951		
103,894,886	ADH4	alcohol dehydrogenase 4 (class II), pi polypeptide [Source:HGNC Symbol;Acc:252]
103,933,564	ADH5	alcohol dehydrogenase 5 (class III), chi polypeptide [Source:HGNC Symbol;Acc:253]
103,959,846	METAP1	methionyl aminopeptidase 1 [Source:HGNC Symbol;Acc:15789]
104,074,224	EIF4E	eukaryotic translation initiation factor 4E [Source:HGNC Symbol;Acc:3287]
104,464,495	238o22	
104,562,652	27o17	
104,591,513	C4orf37	chromosome 4 open reading frame 37 [Source:HGNC Symbol;Acc:28712]
104,991,139		
105,703,406	SW1551	
106,400,730	SW790	
106,584,411	BMPR1B	bone morphogenetic protein receptor, type IB [Source:HGNC Symbol;Acc:1077]
106,689,479		Bone morphogenetic protein receptor type-1B Precursor (EC 2.7.11.30)(CDw293 antigen) [Source:UniProtKB/Swiss-Prot;Acc:O00238]

107,127,204	PDLIM5	PDZ and LIM domain 5 [Source:HGNC Symbol;Acc:17468]
107,465,930	SMARCAD1	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1 [Source:HGNC Symbol;Acc:18398]
107,524,790	HPGDS	hematopoietic prostaglandin D synthase [Source:HGNC Symbol;Acc:17890]
107,828,794		
107,872,512	ATOH1	atonal homolog 1 (Drosophila) [Source:HGNC Symbol;Acc:797]
107,925,517	GRID2	glutamate receptor, ionotropic, delta 2 [Source:HGNC Symbol;Acc:4576]
108,083,510		
108,172,811		
108,876,318		
108,894,996		
110,686,938	S0782	
110,920,156	MMRN1	multimerin 1 [Source:HGNC Symbol;Acc:7178]
111,089,417		
111,179,425		Alpha-synuclein Fragment [Source:UniProtKB/TrEMBL;Acc:Q4PNS0]
111,523,763		
111,724,020	GPRIN3	GPRIN family member 3 [Source:HGNC Symbol;Acc:27733]
111,994,156	FAM13A	family with sequence similarity 13, member A [Source:HGNC Symbol;Acc:19367]
112,041,032		
112,057,551	NAP1L5	nucleosome assembly protein 1-like 5 [Source:HGNC Symbol;Acc:19968]
112,146,916		Probable E3 ubiquitin-protein ligase HERC5 (EC 6.3.2.-)(HECT domain and RCC1-like domain-containing protein 5)(Cyclin-E-binding protein 1) [Source:UniProtKB/Swiss-Prot;Acc:Q9UII4]
112,166,286	PIGY	phosphatidylinositol glycan anchor biosynthesis, class Y [Source:HGNC Symbol;Acc:28213]
112,274,297		
112,296,288		ATP-binding cassette sub-family G member 2 (Brain multidrug resistance protein)(CD338 antigen) [Source:UniProtKB/Swiss-Prot;Acc:Q8MIB3]

112,353,399		
112,366,445		ATP-binding cassette protein ABCG2 Fragment [Source:UniProtKB/TrEMBL;Acc:Q6QAS3]
112,396,255	SPP1	secreted phosphoprotein 1 [Source:HGNC Symbol;Acc:11255]
112,396,677	SPP1	
112,520,831	PKD2	polycystic kidney disease 2 (autosomal dominant) [Source:HGNC Symbol;Acc:9009]
112,702,381	MEPE	matrix extracellular phosphoglycoprotein [Source:HGNC Symbol;Acc:13361]
112,723,776	IBSP	integrin-binding sialoprotein [Source:HGNC Symbol;Acc:5341]
112,723,776	IBSP	
112,786,801	DMP1	dentin matrix acidic phosphoprotein 1 [Source:HGNC Symbol;Acc:2932]
112,817,595	DSPP	dentin sialophosphoprotein [Source:HGNC Symbol;Acc:3054]
112,884,827	SPARCL1	SPARC-like 1 (hevin) [Source:HGNC Symbol;Acc:11220]
112,940,456	NUDT9	nudix (nucleoside diphosphate linked moiety X)-type motif 9 [Source:HGNC Symbol;Acc:8056]
113,057,608	AFF1	AF4/FMR2 family, member 1 [Source:HGNC Symbol;Acc:7135]
113,121,323	HSD17B13	hydroxysteroid (17-beta) dehydrogenase 13 [Source:HGNC Symbol;Acc:18685]
113,146,717	KLHL8	kelch-like 8 (Drosophila) [Source:HGNC Symbol;Acc:18644]
113,417,757		
113,837,962	MAPK10	mitogen-activated protein kinase 10 [Source:HGNC Symbol;Acc:6872]
114,105,524	ARHGAP24	Rho GTPase activating protein 24 [Source:HGNC Symbol;Acc:25361]
114,851,447	SW1980	
114,890,629		
115,191,877	CDS1	CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 1 [Source:HGNC Symbol;Acc:1800]
115,416,373	NKX6-1	NK6 homeobox 1 [Source:HGNC Symbol;Acc:7839]
116,033,227	AGPAT9	1-acylglycerol-3-phosphate O-acyltransferase 9 [Source:HGNC Symbol;Acc:28157]
116,120,483	FAM175A	family with sequence similarity 175, member A [Source:HGNC Symbol;Acc:25829]
116,137,940	MRPS18C	mitochondrial ribosomal protein S18C [Source:HGNC Symbol;Acc:16633]
116,144,462	HELQ	helicase, POLQ-like [Source:HGNC Symbol;Acc:18536]

116,154,763	443f10	
116,198,637	HPSE	heparanase [Source:HGNC Symbol;Acc:5164]
116,245,643	COQ2	coenzyme Q2 homolog, prenyltransferase (yeast) [Source:HGNC Symbol;Acc:25223]
116,323,745		
116,364,509	PLAC8	placenta-specific 8 [Source:HGNC Symbol;Acc:19254]
116,394,181	COPS4	COP9 constitutive photomorphogenic homolog subunit 4 (Arabidopsis) [Source:HGNC Symbol;Acc:16702]
116,475,182	LIN54	lin-54 homolog (C. elegans) [Source:HGNC Symbol;Acc:25397]
116,545,784	THAP9	THAP domain containing 9 [Source:HGNC Symbol;Acc:23192]
116,586,687	SEC31A	SEC31 homolog A (S. cerevisiae) [Source:HGNC Symbol;Acc:17052]
116,649,525	SCD5	stearoyl-CoA desaturase 5 [Source:HGNC Symbol;Acc:21088]
116,839,422	ENOPH1	enolase-phosphatase 1 [Source:HGNC Symbol;Acc:24599]
116,864,757	HNRPDL	heterogeneous nuclear ribonucleoprotein D-like [Source:HGNC Symbol;Acc:5037]
116,895,472	RPLP0	ribosomal protein, large, P0 [Source:HGNC Symbol;Acc:10371]
116,910,628	HNRNPD	heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa) [Source:HGNC Symbol;Acc:5036]
117,779,537	PRKG2	protein kinase, cGMP-dependent, type II [Source:HGNC Symbol;Acc:9416]
117,920,344	BMP3	bone morphogenetic protein 3 [Source:HGNC Symbol;Acc:1070]
118,046,935	C4orf22	chromosome 4 open reading frame 22 [Source:HGNC Symbol;Acc:28554]
118,613,466	FGF5	fibroblast growth factor 5 [Source:HGNC Symbol;Acc:3683]
118,676,477		
118,701,381	PRDM8	PR domain containing 8 [Source:HGNC Symbol;Acc:13993]
118,797,760	ANTXR2	anthrax toxin receptor 2 [Source:HGNC Symbol;Acc:21732]
119,218,036	GK2	glycerol kinase 2 [Source:HGNC Symbol;Acc:4291]
119,272,100	NAA11	N(alpha)-acetyltransferase 11, NatA catalytic subunit [Source:HGNC Symbol;Acc:28125]
119,603,671	PAQR3	progestin and adipoQ receptor family member III [Source:HGNC Symbol;Acc:30130]

119,632,623		
119,735,253	CSN2	casein beta [Source:HGNC Symbol;Acc:2447]
119,751,674	CSN1S1	casein alpha s1 [Source:HGNC Symbol;Acc:2445]
119,824,493		Alpha-S2-casein Precursor [Source:UniProtKB/Swiss-Prot;Acc:P39036]
119,846,554	C4orf40	chromosome 4 open reading frame 40 [Source:HGNC Symbol;Acc:33193]
119,869,518		
119,924,963	ODAM	odontogenic, ameloblast asssociated [Source:HGNC Symbol;Acc:26043]
119,975,507	CSN3	casein kappa [Source:HGNC Symbol;Acc:2446]

Appendix 3

Table of 'positional' candidate genes within 3 Mbp either side of significant SNPs

Trait(s)	Associated SNPs	Genome location[1]	Genes[2]
LS, TBA	ALGA0103270, ASGA0099069	SSC8: 8.6 Mbp	BOD1L, CPEB2, C1QTNF7, CC2D2A, FBXL5, FGFBP1, CD38, BST1, PROM1, FAM184B, LCORL, SLIT2
TBD	MARC0052517	SSC17: 16.8 Mbp	GPCPD1, C20orf196, CHGB, TRMT6, MCM8, CRLS1, LRRN4, PLCB1, PAK7, ANKRD5, MKKS, C20orf94, JAG1
NMUM transformed	H3GA0055446	SSC2: 15.3 Mbp	OR4S2, OR4C45, RPL39, FNBP4, NUP160, AGBL2, PTPMT1, KBTBD4, NDUFS3, FAM180B, C1QTNF4, MTCH2, CELF1, MADD, MYBPC3, SP1, SLC39A13, PSMC3, RAPSN, NR1H3, ACP2, DDB2, PACSIN3, ARFGAP2, C11orf49, LRP4, ARHGAP1, ATG13, HARBI1, DGKZ, MDK, CHRM4, AMBRA1, CREB3L1, PHF21A, GYLTL1B, PEX16, C11orf94, MAPK8IP1, CRY2, SLC35C1, CHST1, SYT13, PRDM11, TP53I11, TSPAN18, CD82, ALX4, EXT2, ACCS, ACCSL, TTC17, API5
NMUM	ASGA0019012	SSC4: 20.9 Mbp	SNTB1, MTBP, MRPL13, COL14A1, DEPDC6, DSCC1, TAF2, ENPP2, NOV, SAMD12, RPS2. EXT1, C8orf85, EIF3H, UTP23, TRTPS1
NMUM	ALGA0121141	SSC6: 82.3 Mbp	RNF125, RNF138,

			FAM59A, MEP1B, KLHL14, C18orf34, ASXL3, NOL4, MAPRE2, ZNF397, ZCAN30, ZNF252, INO80C, GALNT1, MOCOS, GRID2IP, CELF4
NMUM (NMUM, transformed)	ALGA0084499	SSC15: 24.9 Mbp	DDX18, CCDC93, EN1, MARCO, C1QL2, C2orf76, DBI, ERCC3, GYPC
NMUM transformed	ALGA0084780	SSC15: 29.6 Mbp	TSN, TFCP2L1, GLI2, IMP4, PTPN18, TUBGCP5, CCDC115, SCTR, TMEM177, PTPN4, EPB41L5, DLGAP2, MYOM2
NSB transformed	BGIS0004826	SSC14: 57.3 Mbp	ZP4, RYR2, MTR, ACTN2, HEATR1, LGALS8, EDARADD, ERO1LB, GPR137B, NID1, LYST, GNG4, B3GALNT2, TBCE, GGPS1, ARID4B, RBM34, TOMM20, IRF2BP2, TARBP1, C1orf31, SLC35F3, KCNK1, PCNXL2, C1orf57, SIPA1L2
Gestation length	ALGA0021148	SSC3: 107.7 Mbp	CENPA, C2orf18, KCNK3, CIB4, C2orf70, OTOF, HADHB, GPR113, HADHA, FAM59B, RAB10, KIF3C, ASXL2, DTNB, DNAJC27, ADCY3, CENPO, C2orf79, NCOA1, C2orf84, ITSN2, PFN4, TP53I3, FKBP1B, C2orf44, MFSD2B, KLHL29, HS1BP3, RHOB, PUM2, SDC1, LAPTM4A, MATN3, WDR35, TTC32
Gestation length	H3GA0055694	SSC15: 29.1 Mbp	TSN, TFCP2L1, GLI2, IMP4, PTPN18, TUBGCP5, CCDC115, SCTR, TMEM177,

			PTPN4, EPB41L5, DLGAP2, MYOM2
Lactation length	ALGA0004694, MARC0015922	SSC1: 86.75 – 87.5 Mbp	FRK, NT5DC1, COL10A1, DSE, BET3L, FAM26E, FAM26F, FAM26D, RWDD1, LGSN, RIPPLY2, PRSS35, SANP91, ME1, RWDD2A, PGM3, DOPEY1, TPBG, FAM46A
Services, transformed	MARC0036115, MARC0082152	SSC3: 63.4 – 68.0 Mbp	LRRTM4, FAM176A, HK2, SEMA4F, C2orf65, DOK1, LOXL3, HTRA2, AUP1, DQX1, WBP1, TTC31, PCGF1. LBX2, CCDC142, RTKN, WDR54, MOBKL1B, BOLA3, TET3, DGUOK, ALMS1, NAT8, TPRKB, DUSP11. C2orf78, STAMBP, EGR4, FBXO41, CCT7, C2orf7, SMYD5, NOTO, SFXN5, EMX1, SPR, EXOC6B, CYP26B1, DYSF, ATP6V1B1, PAIP2B, NAGK, CD207, CLEC4F, FIGLA, ANKRD53, TGFA, FAM136A, SNRPG, TIA1, PCYOX1, C2orf42, MXD1, GMCL1, ANXA4, AAK1, GFPT1, ANTXR1, GKN1, ARHGAP25, BMP10, APLF, FBXO48, PLEK, CNRIP1, PPP3R1, PNO1, WDR92, C1D, SPRED2, ACTR2, RAB1A, CEP68, SLC1A4
Services, transformed	INRA0015162	SSC4: 82.6 Mbp	RP1, TCEA1, RGS20, ATP6V1H, OPRK1, ST18, PCMTD1, MCM4, CEBPD, KIAA0146, PRRX1, GORAB, METTL11B, KIFAP3, SCYL3, C1orf112, C1orf56, SELE, SELP, F5,

			SLC19A2, C1orf114, ATP1B1, DPT,
Matings	MARC0092197	SSC9: 119.7 Mbp	LAMC1, NMNAT2, SMG7, NCF2, ARPC5, APOBEC4, GLT25D2, RGL1, TSEN15, C1orf21, EDEM3, FAM129A, RNF2, C1orf26, IVNS1ABP, TPR, C1orf27, PDC, PTGS2, CENPF, PTPN14, SMYD2, PROX1, APOA1BP
Matings	ASGA0047195	SSC10: 24.7 Mbp	DENND1B, C1orf53, NEK7, ATP6V1G3, PTPRC, NR5A2, ZNF281, KIF14, IPO9, SHISA4, LMOD1, TIMM17A, RNPEP, ELF3, UBE2T, PPP1R12B, SYT2, KDM5B, CYB5R1, ADIPOR1, KLHL12, RABIF, HABP4, SLC35D2, HSD17B3, C1orf102, FANCC, C9orf3, FBP1, FBP2, DAPK1, GAS1,

1. Approximate genome location(s) for most significant SNP(s)
2. Gene symbols for genes located within 3 Mbp either side of associated SNP. Genes for which the gene symbol is underlined, are involved in multiple embryonic / developmental processes as indicated by the relevant Gene Ontology process terms.

Appendix 4

Tables summarising the information available for gilts/sows from which tissues were collected. Highlighted in yellow are the normal-sized foetus weights and in red the smallest foetuses are indicated.

Roslin Sample

Ids of the pigs		**574**	**509**	**Y24**	**W12**	**Y22**	**W2**	**W8**	**Y26**	**W7**
Born date		27/03/2008	01/03/2008	24/03/2008	27/03/2008	24/03/2008	21/03/2008	24/03/2008	27/03/2008	24/03/2008
Served date		28/10/2008	8/01/2009	13/03/2009	13/03/2009	14/03/2009	15/03/2009	16/03/2009	17/03/2009	17/03/2009
Sire ID		303	303	303	303	303	303	303	303	303
Dam ID		2194	309	2206	2193	2160	2191	2206	2194	2206
Boar ID		303	307	307	LW	303	307	303	LW	307
Parity number		1	1	2	1	1	2	2	2	2
Day of Slaughtering		08/12/2008	18/02/2009	27/04/2009	28/04/2009	27/04/2009	28/04/2009	28/04/2009	27/04/2009	28/04/2009
Stage of pregnancy (days)		42	41	45	46	44	44	43	41	42
Left Horn	**L1**	13.02	12.00	22.40	23.70	16.19	17.13	13.41	6.09	12.68
	L2	10.70	12.00	22.55	22.96	17.10	15.38	14.30	8.69	12.89
	L3	14.87	13.00	22.00	20.89	18.10	18.01	11.97	12.65	13.40
	L4	13.60	13.00	21.90	21.20	17.27	17.72	13.68	12.05	14.52
	L5	13.45	12.00	22.83	25.03	15.06	14.60	11.32	10.27	13.95
	L6	13.08	11.00	21.55	24.18	15.64	17.80	13.65	-	14.13
	L7	3.28	-	20.72	24.22	14.88	19.00	16.15	-	-
	L8	14.90	-	20.01	22.64	19.84	17.31	14.04	-	-
	L9	13.85	-	-	-	16.40	13.10	-	-	-

	L10	-	-	-	-	-	14.42	-	-	-
Right Horn	**R1**	13.65	12.00	21.82	22.94	16.69	16.60	16.77	13.11	12.71
	R2	14.63	11.00	23.08	21.24	17.10	17.49	13.36	11.90	13.95
	R3	15.05	14.00	23.64	19.58	17.18	17.29	13.49	13.25	13.85
	R4	15.71	12.00	25.14	25.24	16.42	16.54	13.00	13.93	14.39
	R5	15.03	12.00	19.06	23.20	-	17.44	14.38	11.18	17.57
	R6	14.06	10.00	22.43	24.02	-	18.53	13.02	-	14.84
	R7	-	-	23.08	22.00	-	18.01	13.16	-	-
	R8	-	-	20.60	22.75	-	15.98	14.48	-	-
	R9	-	-	-	-	-	-	10.83	-	-
Litter size		15	12	16	16	13	18	17	10	12

France sample

ID Gilt		94879	94897	94953	94956	95022	95531	95535	95537	95580
Breed		LW	LW	LW	LW	LW	MS	MS	MS	MS
Born		25/08/2009	26/08/2009	27/08/2009	27/08/2009	28/08/2009	11/10/2009	11/10/2009	11/10/2009	26/10/2009
Sire		1844	1482	1803	1803	2976	4859	4859	04859	3778
Dam		3506	1386	2864	2864	3527	4833	4833	4833	3457
Boar		6702	4874	6702	4874	4874	F4288	F4288	4288	3778
Serving date		27/04/2010	27/04/2010	27/04/2010	27/04/2010	27/04/2010	28/04/2010	28/04/2010	27/04/2010	28/04/2010
D of slaughtering		08/06/2010	08/06/2010	08/06/2010	08/06/2010	08/06/2010	08/06/2010	08/06/2010	08/06/2010	08/06/2010
Age at slaught.		297 days	296 days	295 days	295 days	294 days	240 days	240 days	240 days	225 days
Weight at slaught.		172	163	170	176	132	102	105	101	85
OR left		18	13	9	14	11	13	6	14	7
OR right		3	16	13	8	7	10	17	9	9
OR Total		21	29	22	22	18	23	23	23	16
Litter		42	42	42	42	42	41	41	42	41
Left Horn	**L1**	14.00	12.98	12.03	12.92	14.44	10.26	10.55	11.11	10.18
	L2	14.23	13.36	11.85	11.59	13.88	10.15	11.48	11.30	10.61
	L3	14.37	12.65	14.18	11.57	14.34	10.59	8.84	11.10	9.96
	L4	14.33	14.04	12.95	12.13	12.71	9.98	10.32	10.92	9.81
	L5	15.16		9.98	11.57	12.89	10.76	11.62	12.06	10.89
	L6	14.50		12.59	10.89	14.33	9.84	10.47	8.35	9.89
	L7	11.30		13.45	11.43	12.84	11.06	8.43		8.37

	L8			14.61	13.73	12.60				9.79
	L9			13.63						9.19
	L10			13.14						
Right Horn	**R1**	15.42	12.73	12.73	13.71	12.61	11.18	10.71	11.89	10.10
	R2	15.78	12.72	9.40	12.41	12.00	10.01	11.18	11.53	11.20
	R3	15.06	14.07	12.42	13.07	14.60	10.19	9.77	11.51	11.10
	R4	15.17	12.48	12.74	12.87	13.14	10.46	11.55	11.08	10.73
	R5	15.53	14.21	13.34	13.43	11.77	9.53	10.16	12.16	10.42
	R6			12.78		10.19	10.20	11.56	11.35	10.73
	R7			12.37		12.54	10.68	10.49	12.01	
	R8			10.43		12.48	8.79	9.54	10.43	
	R9			11.74			8.90	9.64	9.65	
	R10			13.81			10.55	9.01		
Litter size		12	9	20	13	16	17	17	15	15