UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Ciència Animal i dels Aliments
Facultat de Veterinària

"Deciphering the genetic architecture of prolificacy related traits in an experimental Iberian x Meishan F_2 intercross"

Ingrid Balcells Ortega

PhD Thesis
June, 2012

Supervisors:
Armand Sánchez and Anna Tomás

El Dr. **Armand Sánchez Bonastre**, catedràtic del Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de Barcelona

i

la Dra. **Anna Tomás Sangenís**, investigadora en la Fundació d'Investigació Sanitària de les Illes Balears de Mallorca

CERTIFIQUEN:

Que l'Ingrid Balcells Ortega ha realitzat sota la seva direcció el treball de recerca "Deciphering the genetic architecture of prolificacy related traits in an experimental Iberian x Meishan F_2 intercross" per a obtenir el grau de doctora per la Universitat Autònoma de Barcelona.

Que aquest treball s'ha dut a terme al Departament de Ciència Animal i dels Aliments de la Facultat de Veterinària de la Universitat Autònoma de Barcelona.

Bellaterra, 11 de Maig de 2012

Dr. Armand Sánchez Bonastre

Dra. Anna Tomás Sangenís

Durant la realització d'aquesta tesi, han sigut moltes les persones que m'han acompanyat, tan a nivell professional com personal. Totes elles han aportat el seu granet de sorra per a que aquest projecte hagi tirat endavant i han fet que pugui recordar aquesta etapa amb un gran somriure a la cara.

Als meus directors de tesi, el doctor Armand Sánchez i la doctora Anna Tomás. Per tota la confiança que heu dipositat en mi, per tots els coneixements que m'heu transmès, per donar-me copets a l'esquena en els moments de més desànim (sobretot en aquests últims mesos) i per mil coses més. Seria una llista molt llarga tot el que us voldria agrair. De tot cor, mil GRÀCIES per tot! I, a l'Anna, vull agrair-li també la seva gran paciència i les seves correccions tant en els articles com en la tesi.

A la doctora Anna Castelló, sens dubte la persona qui més ha contribuït en la meva formació científica, sobretot a nivell de laboratori. Gràcies per estar sempre disposada a donar un cop de mà i pels teus bons consells en tot moment. Llàstima que siguis de Cervera!

Als doctors Sebastian Ramós, Josep Mª Folch, Jordi Jordana, Miguel Pérez-Enciso, Marcel Amills i Jesús Piedrafita per haver contribuït amb la meva formació com a científica amb els vostres cursos de màster i xerrades a peu de passadís.

A totes les nenes, i ara també nen, del SVGM i VETGENOMICS. A l'Olga, a la Laura i a l'Elisenda per les bones estones compartides i pels seus pràctics consells, a l'Anna Mercadé per oferir sempre la seva ajuda, al David per les llargues converses dins i fora el laboratori i a la Griselda, la Rebeca, la Lorena i la Carme per transmetre sempre alegria i felicitat.

Als meus companys que han passat a ser grans amics, al Xavi (els biotecnòlegs primer!), a l'Oriol, a la Natàlia i a la Vero. Són molts els bons moments que hem passat junts (al laboratori, de congressos, viatges, sopars, birres,...) i de tots m'enduc un gran record i que espero que continuïn. I, a l'Oriol, per haver compartit les alegries i les penes del miRNAs. Amb tu, el treball amb RNA ha estat més fàcil!

Als companys que ja no hi són però que han deixat la seva petjada en la meva carrera científica: a l'Anna O., a la Maribel, a l'Oscar, al Jordi E., a l'Abid, a la María S., a la Neus, al John, a la Irene i al Yang Bin.

A tota la resta de companys (espero no deixar-me a ningú, si ho faig, és per culpa les presses!): Alí (per la teva amistat i per les converses interessants que hem mantingut), Arianna (per les divertides converses en els dinars), Anna E. (pels bons moments i per oferir-me sempre la teva ajuda), al Quim C. (per resoldre tots els meus dubtes en estadística, que no han sigut pocs!), a la Marimar (per aguantar-nos), a la Maria B. (per transmetre ànims i pel teu entusiasme), a la Betlem (per aportar fil musical i alegria en tots els moments), a la Marta (per les converses estadístiques), a la Cecília (pels teus somriures) i, a les noves generacions: a la Sarai, a l'Anna P., al Jordi C., al Yuliaxis, a l'Erika, a la Carola i al William, pel vostre suport i amistat, molta sort en les vostres tesis!

A la Susanna Cirera, per haver-me acollit tan bé durant la meva estada a Copenhaguen i haver fet que aquesta hagi estat molt enriquidora per mi, tant a nivell professional com personal. I would like to acknowledge Agnieszka Podoloska to be friendly with me and teach me how Denmark social life is. Moreover, thanks to everybody from the Department of Basic Animal and Veterinary Science (University of Copenhagen) to help me during my stay there.

A tots els investigadors del IRTA-UdL, i de l'INIA per les llargues sessions a l'escorxador i per la vostra col·laboració en el projecte de la meva tesi.

A tota la meva colla targarina, pels molts bons moments viscuts junts, per la vostra incondicional amistat i per fer-me saber que puc comptar amb vosaltres. Sou genials!! També a tots els amics biotecnòlegs: a la Cinta, a la Jone, a l'Alba, al Roger, al Jofre, al Joan, al Pep, a la Bàrbara, a la Paula, a l'Alberto, a la Sara, al Santi, al Xavi, ... per compartir tants bons moments durant i després la carrera.

A l'Eloi, per estar al meu costat en tot moment, per donar-me grans dosis d'ànims en tots els moments de crisi i per ensenya'm a creure més en mi!! Gràcies per TOT!!!

A tota la meva família i, a tu iaia, per transmetre'm el teu amor i per rebre'm sempre amb una gran alegria i una gran abraçada.

Als meus pares, al Jaume i a l'Anna, per estar SEMPRE al meu costat i per haver fet de mi la persona que sóc avui. GRÀCIES!! També al Jaume, el meu germà, per la gran estima i amistat que ens uneix i, perquè tot i que ens fem grans, encara em deixa fer de germana gran.

DECIPHERING THE GENETIC ARCHITECTURE OF PROLIFICACY RELATED TRAITS IN AN EXPERIMENTAL IBERIAN X MEISHAN F_2 INTERCROSS

Reproductive traits are of great interest in the swine industry to improve the pig efficiency production. In a previous study, several QTL affecting reproductive traits were identified in an Iberian (Ib) x Meishan (Me) F_2 population (Noguera *et al.*, 2009, Fernandez-Rodriguez *et al.*, 2010, Rodriguez *et al.*, 2005). In particular, two QTL for the number of piglets born alive (NBA) and the total number of piglets born (TNB) were located on porcine chromosomes 13 (SSC13) and SSC17 (Noguera *et al.*, 2009). To identify genes responsible for the prolificacy QTL on SSC13, four candidate genes (*ITIH1*, *ITIH3*, *ITIH4* and *MUC4*) were analysed. Analyses for *ITIH* gene cluster showed that these genes had an effect on NBA independent of litter size QTL. Results for *MUC4* gene, located within the confidence interval of the QTL on SSC13, determined that a SNP within the *MUC4* gene was associated with NBA and TNB although the effect was stronger for NBA. In addition, uterine *MUC4* expression was two-fold higher in high prolificacy sows. To better understand the genetic basis of prolificacy related traits, transcriptome analyses, which included the study of uterine gene expression as well as the miRNA expression profile in the uterus and in the ovary, was performed. For this, IbxMe F_2 sows displaying extreme phenotypes regarding the prolificacy levels defined as the number of embryos (NE) attached to the uterus at day 30-32 of the gestation were used. In uterus of high prolificacy sows, 101 genes involved in the inflammatory response to stimulus and muscle tissue development were upregulated whereas 196 genes that participated in muscle tissue development, cell junction organization and adhesion, biological regulation, muscle and circulatory system processes, and transport were downregulated. The microRNAome identified the miR-125b-5p, miR-200c-3p, miR-23b-3p, miR-23a-3p and miR-99a-5p as the most abundant miRNAs in uterus of pregnant sows while expression of miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p was associated with prolificacy levels. Among predicted gene target for uterine prolificacy-related miRNAs, 32 were located within the confidence interval of the QTL for NBA and TNB and therefore, they are proposed to be good candidate genes to be further investigated. Importantly, among these candidate genes, it is found *MUC4* gene which is of great interest to be the responsible for the prolificacy QTL on SSC13 because it fulfilled several criteria: it is located within the QTL confidence interval with an effect on NBA and TNB, its expression level varies regarding the prolificacy level of sows, and it is targeted by miR-150-5p, a miRNA that was also found diferentially expressed regarding the prolificacy levels. On the other hand, ovarian miR-146a-5p and miR-142-3p, involved in immune system processes and cellular homeostasis, were differentially expressed regarding prolificacy levels. Four predicted target genes (*LRRK1*, *CCL8*, *CPEB2* and *BAT1*), located within confidence intervals for prolificacy QTL, are good candidate genes to be studied for QTL on litter size. Alternatively, we have designed a new RT-qPCR methodology, by using DNA primers to measure miRNA expression, which is highly specific, sensitive and accurate.

ESTUDI DE L'ARQUITECTURA GENÈTICA DELS CARÀCTERS REPRODUCTIUS RELACIONATS AMB LA PROLIFICITAT EN UN CREUAMENT EXPERIMENTAL F_2 IBÈRIC X MEISHAN

Els caràcters reproductius són de gran interès en la indústria porcina per tal de millorar l'eficiència productiva. En estudis previs, utilitzant un creuament experimental F_2 entre les races Ibèric (Ib) i Meishan (Me) es van identificar varis QTL afectant varis caràcters reproductius (Noguera *et al.*, 2009, Fernandez-Rodriguez *et al.*, 2010, Rodriguez *et al.*, 2005). En particular, es van identificar 2 QTL amb efecte al nombre de garrins nascuts vius (NV) i al nombre total de garrins nascuts (NT) localitzats en els cromosomes porcins 13 (SSC13) i SSC17 (Noguera *et al.*, 2009). Per tal de poder identificar els gens responsables dels QTL en el SSC13, es van analitzar quatre gens candidats (*ITIH1*, *ITIH3*, *ITIH4* and *MUC4*). La caracterització del clúster de gens *ITIH* va permetre identificar que aquests tenen un efecte sobre el NV però que és independent dels QTL associats a la mida de la ventrada. Els anàlisis del gen *MUC4*, que es troba localitzat dins l'interval de confiança del QTL en el SSC13, van determinar una associació significativa entre un SNP dins d'aquest gen i els caràcters NV i NT, tot i que l'efecte era superior pel NV. A més, l'expressió del gen *MUC4* en l'úter és dues vegades superior en truges d'alta prolificitat. Per tal de millorar el nostre coneixement envers l'arquitectura genètica dels caràcters relacionats amb la prolificitat, es va analitzar el transcriptoma a nivell d'expressió gènica en úter així com també els nivells d'expressió de miRNAs tant en úter com en ovari. Aquests estudis es van realitzar utilitzant truges F_2 IbxMe que presentaven fenotips extrems pels nivells de prolificitat definits com el nombre d'embrions (NE) units a l'úter a dia 30-32 de la gestació. En l'úter de les truges d'alta prolificitat es van identificar 101 gens (upregulated) implicats en la resposta inflamatòria enfront als estímuls i en el desenvolupament del teixit muscular. Per altra banda, 196 gens (downregulated) es van relacionar amb el desenvolupament del teixit muscular, l'organització d'unió cel·lular, en processos d'adhesió, en la regulació biològica, en processos del sistema muscular i circulatori i en el transport. L'estudi del microRNAoma va identificar els miR-125b-5p, miR-200C-3p, miR-23b-3p, miR-23-3p i miR-99-5p com els més abundants en l'úter de les truges gestants, mentre que l'expressió de miR-139- 5p, miR-150-5p, miR-27-3p i miR-20-5P es van associar amb els nivells de prolificitat. Entre tots els possibles gens diana per als miRNAs relacionats amb la prolificitat, es troben 32 gens localitzats dins l'interval de confiança per als QTL de prolificitat i, per tant, es van proposar com a bons gens candidats per a ser estudiats. Entre aquests, es troba el gen *MUC4* que és de gran interès per a ser el responsable del QTL en el SSC13 ja que reuneix varis criteris; es localitza dins l'interval de confiança del QTL en el SSC13, té un efecte sobre la mida de la ventrada, el seu nivell d'expressió es relaciona amb els nivells de prolificitat i pot ser regulat pel miR-150-5p, que també es va trobar diferencialment expressat en relació amb els nivells de prolificitat. D'altra banda, en ovari, els miR-146a-5p i miR-142-3p, involucrats en

processos del sistema immunològic i en la homeòstasis cel·lular, estan diferencialment expressats en relació amb els nivells de prolificitat. Quatre gens diana per aquests miRNAs (*LRRK1*, *CCL8*, *CPEB2* and *BAT1*) es troben dins l'interval de confiança per als QTL amb efecte per la prolificitat, fet que fa que siguin bons candidats a ser estudiats. Finalment, es va dissenyar un nou mètode RT-qPCR molt específic, sensible i precís per tal de mesurar els nivell d'expressió dels miRNAs mitjançant l'ús d'encebadors d'ADN.

1. INTRODUCTION

1.1. PIG PRODUCTION

The pig (*Sus Scrofa*) is an animal of great interest for humans (*Homo Sapiens*) that was domesticated at least 9000 years ago (Larson *et al.*, 2007). The permanent source of protein and the excellent organoleptic characteristics of its meat were the main reasons for their domestication (Jones, 1998). Furthermore, pigs were used to obtain clothes and shoes from their skin, and tools and weapons from their bones. Nowadays, although in a different way, the porcine industry is still very important since it is the meat with more consumption around the world. Its importance is captured with the pig production representing the 40% of world meat production (FAOSTAT, 2011). Asia is the first porcine meat producer (57%) followed by Europe (25%) (FAOSTAT, 2011) where Germany is the first European porcine producer with 26,900,800 pigs (19%) and Spain is the second one with 25,704,000 pigs (18%) (EUROSTAT, 2011). In Spain, Catalonia concentrates the major porcine census (26%) representing the 41% of the Spanish total porcine meat production (MARM, 2011).

The economical importance of the pig industry is evidenced in Spain. The total pig production represents the 35% of the final livestock production and the 11% of the total agricultural production (MARM, 2011). In Catalonia, the importance of pig production is even greater representing the 58% of the final livestock production and the 34% of the total agricultural production (DAAM, 2010).

1.1.1. Efficiency improvement in pig production

Since the very beginning, the main objective of swine industry has been to produce the greatest improvement per unit of time (Dickerson and Hazel, 1944). So, the efficiency of pig production needs to be increased in order to maximize the economical yield.
Different traits have been the focus for the pig industry. During the early part of the last century, breed characteristics and physical appearance were the most importance traits to select for. Specially, the fat and the manure of pigs were of great interest for farmers (Merks, 2000). Later on, starting in the 1950s, genetic improvement efforts were still focused on production traits. Mainly, efforts were focused on the reduction

of backfat thickness and the improvement in growth rate to increase lean meat production lowering the pork production cost price (Merks, 2000). During the last decades, reproduction traits, like litter size, became to be of great interest in pig industry due to the fact that good fecundity is directly related to sow's productive life (Distl, 2007). More recently, the breeding goals have been adjusted to the society demand focused on non-economically efficient pork production traits such as the welfare and health of pigs, the ecological effects of pork production, and the healthiness and sensory quality of pork (Kanis *et al.*, 2005).

1.2. REPRODUCTIVE TRAITS

1.2.1. Biology of porcine reproduction

The domestic sow is a polytocous specie. Usually, gilts reach the puberty at 6-7 months of age. However, Chinese breeds like Meishan can achieve puberty at an early age of 3-4 months. After puberty, the oestrous cycle is characterised by lasting 18-24 days which includes a proestrous period (~ 2 days), an estrous (2-3 days), a metraestrous (1-2 days) and the rest of the days correspond to diestrous. Ovulation occurs during the second half of the oestrous where 15 to 30 follicles are ovulated. Mating schemes of two services during the oestrous, right before the ovulation occur, lead to very high conception rations being the fertilization ova success rate greater than 95%. Then, the pregnancy lasts from 114 to 116 days with an average of litter size ranging from 4 to 16 piglets at each litter depending on the breed. As it is evidenced, not all the ovulated ova are developed to piglets suggesting a loss of potential piglets during pregnancy. Specifically, around 30-50% of the ova released from the ovary do not survive the gestation process (Pope, 1994).

Prenatal losses during pregnancy have an essential impact on litter size (Pope, 1994). There are two critical phases in the prenatal losses; embryonic or foetal losses, depending on the developmental stage of the conceptus. The first phase includes the implantation process between days 12-18 of pregnancy where the embryonic losses can reach up to 40% with an average of 20-30% (Distl, 2007). At this period, the embryonic losses are mainly due to a non equidistant spacing of embryos throughout

the uterus, a slow trophoblastic elongation, a wrong establishment of conceptus-uterine attachment and a rejection of the immune system by the maternal system (Spotter and Distl, 2006). The second phase occurs during the foetal growing phase, when the embryo is already attached to the endometrium, from day 30 of pregnancy. During this period, foetal losses range between 5% and 50% with an average of 10-20% (Pope, 1994). In this phase, the uterine capacity and the placental development become critical for the growth of fetuses (Spotter and Distl, 2006). Perinatal mortality has also important consequences, with incidences of stillborn piglets ranging from 10% to 15% (Herpin *et al.*, 2001). Piglet death in this period is mainly caused by uterine infections or by asphyxia during parturition due to an interruption of the oxygen flow through the umbilical cord which account for 60-70 % of the perinatal losses (Alonso-Spilsbury *et al.*, 2005, Herpin *et al.*, 2002).

After birth, piglets need to adapt their body to the extrauterine environment. Specially, piglet survival depends mainly on thermoregulation (Curtis, 1974; Alonso-Spilsbury *et al.*, 2005). However, other factors such as neonatal asphyxia, lesser vigor and reduced colostrum intake are also involved in the piglet mortality after birth (Alonso-Spilsbury *et al.*, 2005). Overall, the mortality rate related to reproduction process is rather high in pigs, close to half of the total embryos.

1.2.2. Genetic basis for reproductive traits

Reproductive traits, like most economically interesting traits, have a multifactorial inheritance; they are influenced by many genes as well as by environmental factors. In general, female reproductive traits have low to moderate heritabilities (Table 1.1) indicating that the additive genetic variation of these traits is assessed by genetics and by other environmental factors such as health and nutrition status, which can have a large impact on them. Among the female reproductive traits, the most heritable traits are those depending only on the female genotype such as age of puberty and ovulation rate with heritabilities above 0.3. On the other hand, fertility and prolificacy traits like litter size, piglet survival rates and conception rate, which are the result from complex interactions between the sow, the boar and the embryo/piglet genotypes,

have low heritabilities evidencing difficulties to be improved through conventional selection.

Table 1.1. Heritabilities (h^2) estimation for female reproductive traits.[1]

Trait	Mean h^2	Range
Age at puberty	0.37	0 - 0.73
Ovulation rate	0.32	0.10 - 0.59
Litter weight at birth	0.24	0 - 0.54
Prenatal survival rate	0.15	0 - 0.23
Total number of piglets born	0.11	0 - 0.76
Number of piglets born alive	0.10	0 - 0.66
Conception rate	0.10	0 - 0.29
Number of piglets weaned	0.08	0 - 0.10
Birth to weaning survival rate	0.05	0 - 0.13

[1] Extracted from Bidanel (2011).

Litter related traits are mainly influenced by the sow genes although fecundity is controlled likewise by paternal and maternal genes showing both similar heritabilities (Varona and Noguera, 2001). Maternal effects are higher than piglet effects until the piglet has grown enough that it do not depend on sow's milk, at 3-4 weeks of age (Kaufmann *et al.*, 2000,Solanes *et al.*, 2004,Rosendo *et al.*, 2007a). After weaning, the maternal effects are strongly reduced or even vanished.

The preweaning environment provided by the sows to the litter traits would affect the decreasing of the selection efficiency for these traits. Nevertheless, the maternal effects have been estimated but controversial results have been obtained. Some authors have detected small or non maternal effect (Chen *et al.*, 2003, Perez-Enciso and Gianola, 1992) while others have determined a significant effect (Irgang *et al.*, 1994, See *et al.*, 1993, Ferraz and Johnson, 1993, Southwood and Kennedy, 1990). These controversial results show the difference between pig populations and/or management conditions. In any case, ignoring the maternal effects would reduce the selection response in prolificacy traits (Roehe and Kennedy, 1993).

Phenotypic and genetic correlations between reproduction traits are shown in (Table 1.2). Litter size at birth and at weaning is strongly positively correlated. Conversely, litter size related traits are negatively correlated with piglet survival. Litter size has also negative genetic correlation with average birth weight which is positively correlated

with piglet survival rates. In relation to other economically interesting traits, litter size was thought to be weakly correlated with performance traits (Brien, 1986; Haley *et al.*,1988) although recent studies determined a negative genetic correlation between growth traits and litter size (Tribout and Bidanel, 2008a, Holm *et al.*, 2004, Hermesch, 2000, Hermesch, 2000, Zhang *et al.*, 2000, Ducos and Bidanel, 1996). However, the vast majority of studies determine a slight positive correlation between the carcass lean content trait and litter size (Bidanel, 2011).

Table 1.2. Means of genetic and phenotypic correlations[1] among reproductive traits[2].

	AP[3]	OR	TNB	NBA	NW	BWSR	LBW
AP		- 0.06	- 0.04	0.07	0.09	-	- 0.10
OR	0.05		0.32	0.24	0.01	- 0.38	0.24
TNB	- 0.03	0.13		0.92	0.73	- 0.15	0.62
NBA	- 0.03	0.12	0.91		0.81	- 0.14	0.64
NW	- 0.01	0.03	0.71	0.79		0.55	0.67
BWSR	-	- 0.11	- 0.12	- 0.22	0.15		- 0.07
LBW	- 0.03	0.07	0.79	0.82	0.71	0.09	

[1]Genetic correlations above the diagonal; phenotypic correlation below
[2] Extracted from Bidanel (2011)
[3]AP = Age at puberty; OR = ovulation rate; TNB = Total number born; NBA = Number born alive; NW = number weaned; BWSR = birth to weaning survival rate; LBW = little weight at birth.

1.2.3. Improvement of reproductive traits

Traditionally, genetic improvement programs have used animal selection based on the pedigree information and observable phenotypes assuming that they represent the whole effect of all genes and environment. Thus, for each animal, an estimated breeding value (EBV) was predicted by using the methodology of Best Linear Unbiased Prediction (BLUP), then, the animal with the highest EBV for the breeding objective was selected as reproducer. Although some improvements in reproduction traits have been obtained by using BLUP approach, it has several limitations. One of the main limitations is the low heritabilities of the reproduction traits that limit the accuracy and efficiency of the predicted value. Moreover, reproduction traits are expressed late in life. Thus, the measure of the traits implies high generation intervals in order to select the best animals. Finally, it has to be considered that animal selection is based on families rather than individuals which provoke high inbreeding rates (Bidanel, 2011).

During the eighties, the implementation of molecular genetics, in particular the discovery of new classes of DNA polymorphism, prospects these limitations. In this sense, molecular genetics were implemented with the aim of identifying genes and chromosomal regions affecting traits of great interest in livestock production (Andersson, 2001). Many genetic markers were associated with traits of interest showing their potential use in genetic animal improvement programmes to directly select animals through marker-assisted selection (MAS) (Dekkers and Hospital, 2002). However, the application of MAS in commercial pig breeding has been limited. Potential reasons for their restricted use are the limited number of studies validating the genetic markers, the high cost of genotyping, the limited amount of genetic variation explained by the genetic markers, the fact that the analysis of genetic markers were based on within-family studies and/or experimental intercrosses, the lack of genetic marker results replication and/or their inconsistent results across pig populations and the fact that genetic markers were developed by using wild boar or Chinese breeds and the favourable effects found were fixed in commercial breeds (Bidanel, 2011). The main use of MAS in commercial pig breeding has been related to reduce the incidence of deleterious effects or to increase the frequency of alleles with favourable effects (Bidanel, 2011). A prime example of the use of MAS in porcine crossbreeding programmes falls in the paternally expressed *IGF2* gene which has shown favourable effects on lean growth and reproductive performance owing to an imprinting effect. By this way, sows that are produced from a cross between a boar homozygous for the wild (G) allele for *IGF2* and then are mated with terminal sire homozygous for the mutant (A) allele for leanness, will produce pigs that will be lean because of the sire allele is the lean allele. Moreover, as sows have inherited the wild allele (G) from their sire, they will have increased reproductive performance although they will not pass the increased fatness trait on to their progeny (Dekkers *et al.*, 2001). Recent technology advances which include genome sequencing, the identification of large numbers of genetic markers across the genome in the form of SNPs and the accession to high-throughput genotyping techniques on individual animals with a relatively low price, have led to develop new statistical analyses for molecular data. In this sense, (Meuwissen *et al.*, 2001) proposed a variant of MAS called genomic selection in which genetic markers covering the genome are used in order to obtain a

more accurate predicted value and to remove limitation from MAS (Goddard and Hayes, 2007).

Pig breeders started to take reproduction traits into consideration 15 years ago when they were included in breeding programmes. Most pig producers practise the crossbreeding as a tool to enhance reproductive performance as well as to take advantage of the hybrid vigour or heterosis. The work of Smith (1964) and Moav (1966) establish the basis for crossbreeding through the uses of specialized sire and dam lines. Both lines are specialized for different traits; the dam line is specialized in reproductive traits whereas sire line is specialized in production traits. In order to reach gains in reproduction traits, maternal lines are selected for these traits.

Several selection experiments have been implemented with the aim of increasing litter size. Direct selection for litter size can effectively increase the reproductive performance (Holl and Robison, 2003, Noguera *et al.*, 2002, Lamberson *et al.*, 1991), especially in large populations. Moreover, litter size could also be improved through indirect selection for components of litter size such as the index of ovulation rate and the embryo/prenatal survival (Rosendo *et al.*, 2007b,Ruiz-Flores and Johnson, 2001,Johnson *et al.*, 1984), uterine capacity (Christenson *et al.*, 1987) or placental efficiency (Wilson *et al.*, 1999). Nevertheless, indirect selection could not outperform the direct selection for litter size (Bidanel, 2011). Interestingly, results from successful selection for litter size show no improvements in prenatal survival (Tribout *et al.*, 2003, Driancourt *et al.*, 1992) evidencing the antagonistic relationship between litter size traits and the piglet survival (Bidanel, 2011). In agreement with this, selection for total number born has resulted in an increase of the number of stillborn piglets (Blasco *et al.*, 1995, Ruiz-Flores and Johnson, 2001, Canario *et al.*, 2007b). Over the last years, there has been an increasing interest of pig breeders for decrease piglet mortality. It has been described that survival rates and litter weight are positively correlation. Thus, it is thought that increasing litter weight would increase piglet survival rates (Bidanel, 2011). However, increasing piglet birth weight would be associated with higher nutritional demand from litter during the gestation and lactation that could be harmful for the sow (Bergsma *et al.*, 2008, Canario *et al.*, 2006)

and furthermore, it would increase farrowing difficulties (*Canario et al., 2006*). Other authors have suggested to improving the homogeneity of piglet weight to decrease piglet mortality. Indeed, some studies reported positive relationships between within-litter variation in birth weight and the pre-weaning mortality (Huby *et al.*, 2003, Milligan *et al.*, 2002, Knol *et al.*, 2002, Roehe and Kalm 2000). Additionally, piglet maturity at birth is another important factor related to piglet survival. In fact, it has been determined that piglets from Meishan dams have a higher probability of survival due to they are more mature at birth than European/American breeds (Le Dividich *et al.*, 1991,Le Dividich *et al.*, 1991,Canario *et al.*, 2007a). However, finding the criteria to assess piglet maturity remains a challenge.

1.3. MOLECULAR GENOMIC APPROACHES TO IMPROVE PIG REPRODUCTION

1.3.1. Detection of quantitative trait loci (QTL)

Reproductive traits, like many other important economical production traits, are quantitative traits which have measurable phenotypic variation controlled by several genes and/or influenced by environmental factors. This phenotypic variation usually follows a normal distribution with values that can be discrete (number of piglets born alive) or continuous (weight). An important way to uncover the genetic basis of quantitative phenotypic variation is through the identification and characterization of QTL. A QTL is defined as a region of the genome that contains one or more genes that affect the variation of a quantitative trait (Andersson, 2001).

Since the identification of the first QTL in livestock species in porcine chromosome 4 associated with the growth rate, fatness and the length of the small intestine (Andersson *et al.,* 1994), several QTL experiments have been performed. In pigs, as in other livestock species, the identification of QTL has been commonly done in experimental designs between crossing two breeds highly divergent for the traits of interest. By this way, the maximal statistical power and with the minimal number of individuals can be achieved (Ron and Weller, 2007). An F_1 is generated which is then intercrossed (F_2) or backcrossed (BC) to one of the parental lines. Most of these

experimental crosses have been carried out by mating American/European breeds to Chinese breeds (Georges, 2007).

However, crosses between breeds are generally not realistic compared with the crosses used in pig industry. The porcine breeding programs exploit the genetic variation that exists within breeds instead of the variation between breeds (Georges, 2007). Thus, recent studies have been done using commercial breed intercrosses to validate which QTL obtained from experimental crosses were segregating and/or to identify new QTL.

The identification of QTL is possible since a large number of polymorphic genetic markers are identified. These allow building linkage maps based on recombination frequency which are used for ordering and measuring the distance between two loci on a chromosome. The markers most used are microsatellites, tandem repetitions of two to six nucleotides, and Single Nucleotides Polymorphisms (SNPs), changes of a single base pair. Then, the statistical association between a genetic marker and the quantitative trait of interest can be tested. Statistical significant associations would indicate that a genetic marker is linked to a QTL (they are in linkage disequilibrium). The power and the accuracy of the QTL analyses increase when more genetic markers are used. Recently, high-density linkage maps based on SNP markers have been developed in many livestock species. In pigs, the Illumina PorcineSNP60 BeadChip (Ramos *et al.*, 2009) which contains a total of 62,163 SNPs covering all the genome, has provide an opportunity to carry out whole-genome association studies (WGAS) that would improve the identification of QTL.

There exist different statistical approaches which are implemented in different software for QTL mapping. However, the majority of QTL in livestock species are identify through the interval mapping concept (Lander and Botstein, 1989). By using this, the position and the effect of a putative QTL is estimated every x centiMorgan (cM) based on the information of the flanking markers. Interval mapping method has been implemented in several statistical approaches like least squares regression and maximum likelihood (ML) methods. In the present thesis, QTL analysis was performed using a new interval mapping statistical approach which is based on the ML method

and the mixed model methodology and it has been implemented in Qxpak software (Perez-Enciso and Misztal, 2004). The assumed model is:

$$y = \mu + Xb + Wq + Zu + e \quad \text{(model 1)}$$

where y was the phenotypic data vector for the trait of interest, μ is the overall mean, b are the fixed effects (i.e. year-season, parturition order and environmental effect), q are the QTL effects, u are the polygenic effects, e is the residual term and X,W and Z are the incidence matrix.

The analysis of QTL described above has allowed the identification of a great number of QTL with effects focused on the identification of additive and/or dominant QTL effects. However, it is thought that only a (small) fraction of the genetic variability of traits such as reproductive traits has been identified. In order to detect the non-additive gene effects for these traits, more complex QTL models including imprinting and epistatic interactions effects have developed in order to better adjust the QTL analyses to a real biological context.

1.3.1.1. QTL for litter size in pigs

During the last decades, many studies have performed with the aim of identifying a large number of QTL. By this way, 6,818 QTL have been described representing a total of 585 different traits from 290 publications (PigQTLdb, Release 17). Despite the large amount of QTL, only 322 (4.72%) QTL are affecting reproductive traits (PigQTLdb, Release 17). The limited number of reproductive QTL reflects the complexity of reproductive traits which have low heritabilities and can only be measured in females until later in life.

QTL determined for litter size related traits are summarized in Table 1.3. Interestingly, the number of QTL for litter size is relatively low compared with other reproductive traits and moreover, only a few of them have reached the genome-wide significance level (Noguera *et al.*, 2009, Tribout *et al.*, 2008). Furthermore, there is a lack of overlapping between different QTL studies as it is reflected in the widespread location of them, on 13 different chromosomes (Bidanel, 2011). Most of these QTL analyses

have dissected the additive and the dominant components of litter size. In the majority of QTL analyses for reproductive traits, non-additive effects like epistatic interactions and imprinting have neglected although they are expected to explain a substantial portion of genetic variation of reproductive traits (Carlborg and Haley, 2004).

Table 1.3. QTL for litter size related traits[1].

Trait	Pig chromosome number	Population[2]	Variance (%)	Reference
Total number born (TNB)	11	LW x LR	5.1	(Cassady *et al.*, 2001)
	7, 12, 14, 17	LW/LR x MS	2.7 - 8.8	(de Koning *et al.*, 2001)
	8	LW x MS	-	(King *et al.*, 2003)
	7, 15	DU x ER	2.8 - 4.3	(Li *et al.*, 2009)
	13, 17	IB x MS	-	(Noguera *et al.*, 2009)
	6	YO x MS	-	(Wilkie *et al.*, 1999)
Number born alive (NBA)	11	LW x LR	-	(Cassady *et al.*, 2001)
	6, 15	DU x ER	3.7 - 5	(Li *et al.*, 2009)
	13, 17	IB x MS	-	(Noguera *et al.*, 2009)
	7, 16, 18	LW, LR	-	(Tribout *et al.*, 2008)
Number stillborn (NS)	5, 13	LW x LR	7.9	(Cassady *et al.*, 2001)
	12, 14	LW x LR	-	(Holl *et al.*, 2004)
	7, 8	DU x EU	3.7 - 5	(Li *et al.*, 2009)
	6, 11, 14	LW, LR	-	(Tribout *et al.*, 2008)
	4	YO x MS	-	(Wilkie *et al.*, 1999)

[1]Extracted from Bidanel (2011).
[2]DU = Duroc; ER = Erhualian; IB = Iberian pig; MS = Meishan; LR = Landrace; LW = Large White; YO = Yorkshire.

1.3.1.2. QTL previously identified in the MEIBMAP project

Within the MEIBMAP project, an experimental F_2 intercross was created as it is showed in Figure 1.1. The population was created by mating 18 Meishan (Domaine du Magneraud, INRA, France) sows and 3 Iberian from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) boards which generated an F_1 composed by 8 boars and 97 sows that were intercrossed to obtain the F_2 progeny. Among F_2 sows, 255 F_2 sows were randomly selected for mating which generated a total of 881 parities with an average of 3.45 parities per F_2 sow. During 4 consecutive parities, the total number of piglets born (TNB) and the number of piglets born alive (NBA) was annotated obtaining for each sows with an average phenotypic data of 8.69 (±3.04) and 9.02 (±3.10) for

TNB and NBA, respectively. In the fifth reproductive cycle, sows were slaughtered at 30–32 days of gestation and the number of embryos (NE) and the number of corpus luteum (CL) were recorded.

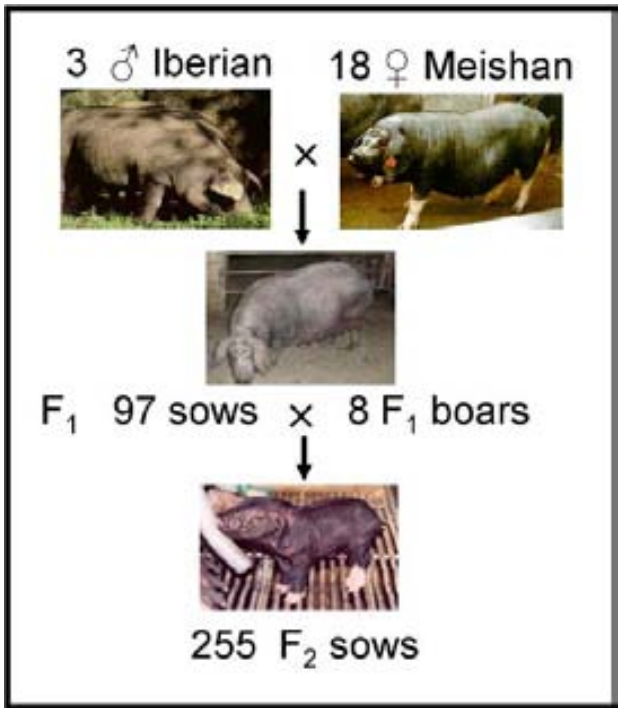


Figure 2.1 Experimental F_2 Iberian x Meishan intercross.

The two parental breed used to create this F_2 intercross are highly divergent European and Asian breeds which present extreme phenotypes for prolificacy traits. The Meishan breed is considered one of the most prolific breeds (average of 14.3 piglets born alive per parity, Bidanel *et al.,* 1989) whereas Iberian breed is considered a very low prolificacy breed (average of 7.0 piglets per parity, Silió *et al.,* 2001). Additionally, these breeds are also highly divergent at the genetic level (Alves *et al.*, 2003).

Table 1.4. Summary of significant QTL at genome wide level analyzed within the MEIBMAP project for reproduction traits.

Trait[a]	SSC[b]	Position cM	Reference
TN	5	29	(Rodriguez *et al.*, 2005)
	10	71	
	12	70	
NBA	13	50	(Noguera *et al.*, 2009)
	17	22	
TNB	13	55	
	17	22	
NBA	12	15	(Fernandez-Rodriguez *et al.*, 2010)
		91	
TNB	12	14	
		91	

[a]TN = Teat Number; NBA = Number of piglets born alive; TNB = Total number of piglets born; BF150 = backfat thickness at 150 days of age.
[b]*Sus Scrofa* chromosome.

The QTL identified within the MEIBMAP project are summarized on Table 1.4. Firstly, three significant QTL at genome wide level associated with teat number trait were identified in SSC5, SSC10 and SSC12 (Rodriguez *et al.*, 2005). Importantly, all these three QTL explained the 30% of the phenotypical variance for this trait. Afterwards, for the first time, two highly significant QTL at genome wide level were identified for the number of piglets born alive (NBA) and total number of piglet born (TNB) in SSC13 and SSC17 (Noguera *et al.*, 2009). The proportion of phenotypic variance explained by both QTL ranges from 2% to 3% for both TNB and NBA traits. For both QTL associated with NBA, highly additive and dominance effects were detected although the direction of these effects depends on the chromosome. For instance, favourable QTL effects on SSC13 come from Meishan breed which is agreement with the phenotypical differences observed in the purebred parental lines. Conversely, favourable QTL effects on SSC17 come from cryptic Iberian allele showing the complexity of the porcine reproduction traits. Furthermore, for NBA and TNB, several epistatic QTL interactions were detected demonstrating that the phenotypic variance of these traits can be highly influenced by a complex network of interacting locis. Overall, a total of 18 epistatic interaction were detected which involved 13 out of the 18 pig autosomes (Figure 1.2) (Noguera *et al.*, 2009). As an important feature was that QTL for NBA on SSC13 and SS17 did not show any significant epistatic QTL and therefore, only additive effect can be related to these QTL. Importantly, in a posterior work, the 2 epistatic QTL

interactions on SSC12 for NBA and TNB traits were further validated by (Fernandez-Rodriguez *et al.*, 2010) showing similar results as previously reported. On the other hand, within MEIBMAP project, QTL for growth and fatness traits were identified. Specifically, on SSC6, two significant QTL in similar positions for backfat thickness and for body weight were identified (Munoz *et al.*, 2009). Afterwards, also in SSC6, a significant QTL was characterized for backfat thickness at 150 days of age (Tomás *et al.*, 2011).

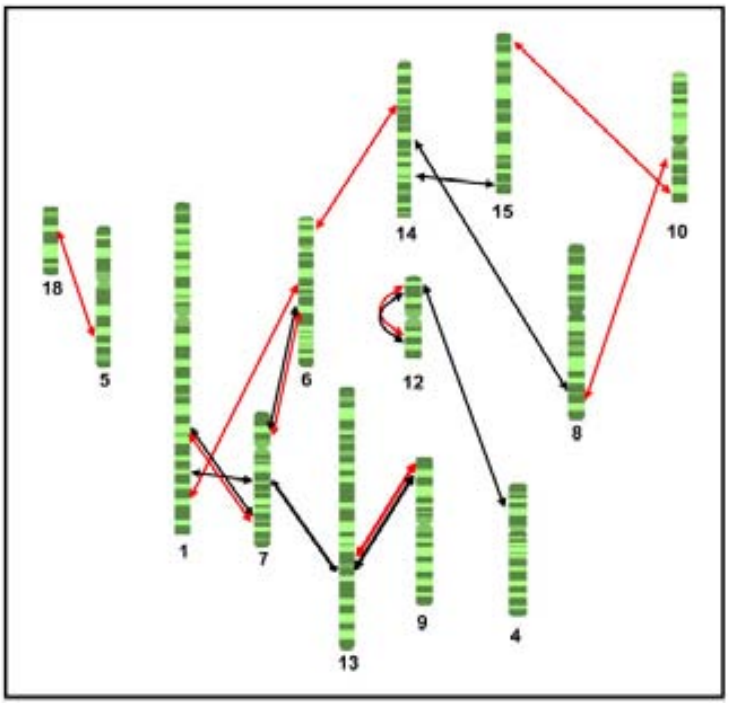


Figure 1.2. Network representation of epistatic QTL interaction for NBA (red arrows) and TNB (black arrows).

1.3.2. Analysis of candidate genes

Dissecting the genetic architecture of complex and quantitative traits can also be performed by a candidate gene approach. Importantly, candidate gene analysis is a complementary strategy to QTL mapping. In fact, it is usually an indispensable procedure after the identification of QTL with the main objective of identifying the true causal mutation also called Quantitative Trait Nucleotide (QTN) to incorporate them in breeding programs through MAS with the main aim of selecting animals based

on the genotype of molecular markers associated with the trait of interest. Candidate gene analysis is likely to be a more effective tool since it has more statistical power than QTL analysis (Tabor *et al.*, 2002). Nevertheless, few studies have successfully discovered the true causal mutation in livestock species (i.e. Clop *et al.,* 2006, Van Laere *et al.,* 2003) although a large number of QTL have been described (PigQTLdb, Release 17). In general, QTL identified have large confidence intervals (typically 20 - 30 cM) which could assemble hundreds of genes. All genes within the confidence interval of the QTL are considered putative for the trait investigated. However, it is not possible to examine all this large amount of genes due to it is time consuming and also expensive. Thus, it is inevitably performing an exhaustive study of all genes in order to select some of them to analyze and to perform association analysis.

There are two main criteria to select candidate genes (Zhu and Zhao, 2007). On one hand, a gene can be chosen as a candidate gene because it has a known physiological role in the phenotype of interest (**physiological candidate genes**). This approach could be performed without any previous QTL analysis but it requires a well-known physiological, biochemical or functional knowledge which is generally finite or sometimes not available at all. Alternatively, the candidate gene could be selected by its genomic location within a described QTL (**positional candidate genes**). In species for which the complete genome sequence is available, the selection of candidates is straightforward. However, in species without a complete genome sequence, identification of positional candidate genes can be performed by comparative mapping. Recently, in pigs, a complete genome sequence (Sscrofa10.2) is available at 24x coverage provided by the International Swine Genome Sequencing Consortium from a Duroc female (GenBank accession number AEMK00000000, September 2011).

The effect of a candidate gene on a quantitative trait is tested typically through association studies between the genotype of the candidate gene and the phenotype of the trait of interest. In livestock species, in F_2 intercrosses as well as in commercial population, the statistic methodology most used is based on the Animal Model which allows testing the effect of the gene by the inclusion of fixed and random effects.

$$Y = Xb + Zu + Za + e \quad \text{(model 2)}$$

A standard Animal Model (model 2) is composed by a *y* vector which contains the phenotypic data for the trait of interest, a *b* vector with the fixed effects (such as age, parturition order and environmental effects), *u* and *a* vectors containing the polygenic and additive genetic effects, respectively, as random effects and *e* vector which includes the random residual term. In *a* vector the number of copies (0, 1, or 2) of the tested allele for each individual is included. X and Z are the incidence matrices that relate observations to *b* and *a* vectors. Ideally, these analyses are conducted within breeds or lines where the LD is extended over short distances. However, in experimental F_2 intercrosses, the LD is very extensive and this could bias the estimation of the effects. The confounding effect of the extensive LD could be excluded by using the maker-assisted association test (MAAT) (model 3) (Zhao *et al.*, 2003).

$$y= Xb + Zu + Wq + Za + e \qquad \text{(model 3)}$$

MAAT (model 3) uses a standard animal model (model 2) introducing the QTL effects (q) into the model. By using MAAT, the additive genetic effect of a gene is tested by taking the QTL effect into consideration. Thus, it can be evaluated if the effect of the gene is by itself or it is due to it is closely to a QTL. Although the false positive results decreases, the power of the test is reduced. However, results from association analyses must be validated in different pig population and breeds which is, in fact, the best way to prove the consistency of the association studies results.

1.3.2.1. Candidate genes for reproductive traits analysed in pigs

The first studied candidate genes for reproductive traits in pigs were selected because their function was relevant in biological processes related to reproduction (mainly focused on litter size) or because disruption of the gene in mice knockout models yielded reproductive defects (Table 1.5). The first candidate gene studied with an association with litter size traits was *Estrogen receptor 1* (*ESR1*) gene (Rothschild *et al.*, 1996) but its confirmation has been controversial in further works showing that the effect of this gene could be different depending on the population (Spotter and Distl, 2006). Other genes like *Prolactin Receptor* (*PRLR*) (Vicent *et al.,* 1998), *follicle*

stimulating hormone beta polypeptide (*FSHβ*) (Li *et al.,* 1998), *leptin receptor* (*LEPR*) ((Chen *et al.*, 2004)) gene and *retinol binding protein 4* (*RBP4*) gene (Rothschild *et al.*, 2000) were investigated and they were successfully associated with litter size traits although their effects also depended on the population studied (Spotter and Distl, 2006).

Afterwards, the implementation of linkage studies for the identification of QTL resulted in the detection of several reproductive QTL in pigs (Table 1.3). Interestingly, some candidate genes selected based on their known reproduction function and with reported effects on litter size traits were located within the confidence interval of any litter size QTL reported. For instance, *LEPR* gene (Chen *et al.*, 2004) was located within QTL on SSC6 (Tribout *et al.*, 2008,Li *et al.*, 2009,Wilkie *et al.*, 1999), *PRLR* gene (Vicent *et al.,* 1998) closely to NBA QTL on SSC16 (Tribout *et al.*, 2008) and *RBP4* gene within TNB QTL on SSC14 (Muñoz *et al.,* 2007). However, not all candidate genes were located within QTL confidence intervals. There is a lacking in concordance between candidate genes and QTL analyses. For example, *ESR1* gene has not been assigned to any described litter size QTL. This lack of correspondence could be due to several reasons (Kirkpatrick, 2002): (1) the modest effect of candidate genes analyzed are undetectable with genome-wide QTL search; (2) a lack of segregation of the candidate genes alleles in some populations; (3) the candidate gene marker may be a linked marker with heterogeneity of linkage phase eliminating association; or (4) the candidate gene effect may be to chance because of the statistical methods used.

Once reproductive QTL were identified, the location of them has been very useful to select candidate genes in order to identify the responsible gene and the causative mutation. Many efforts have been performed to further investigate reproductive QTL by analyzing different candidate genes within confidence intervals of QTL (i.e. (Fernandez-Rodriguez *et al.*, 2010,Du *et al.*, 2009,Buske *et al.*, 2005)) but until now, any causal mutation for reproductive QTL has been described.

Taula 1.5. Summary of candidate genes associated with reproductive traits[1].

Gene	Gene Symbol	SSC[2]	Trait[3]	Reference(s)
Estrogen receptor 1	*ESR1*	1	NBA TNB TN	Rothschild *et al.*, 1996 Short *et al.*, 1997 van Rens *et al.*, 2002V Goliasova and Wolf, 2004 Horogh *et al.*, 2005 Munoz *et al.*, 2007
Follicle stimulating hormone beta	*FSHβ*	2	NBA TNB NW LWW GL	Li *et al.,* 1998 Li *et al.,* 2008
Erythropoietin receptor	*EPOR*	2	UC	Vallet *et al.*, 2005
Leptin receptor	*LEPR*	6*	Litter size	Chen *et al.*, 2004
Fucosyl transferase 1	*FUT1*	6*	NBA TNB	Horak *et al.*, 2005
Ring finger protein 4	*RNF4*	8*	NBA TNB	Niu *et al.*, 2009
Properdin	*BF*	7*	NBA TNB	Buske *et al.*, 2005
Gonadotrophin releasing hormone receptor	*GNRHR*	8*	OR	Jiang *et al.*, 2001
Osteopontin	*OPN*	8*	NBA TNB	Korwin-Kossakowska *et al.,* 2002
Leukaemia inhibitory factor	*LIF*	8	NBA	Spotter *et al.*, 2009
Aldo keto reductase 1C2	*AKR1C2*	10*	AP OR TN	Nonneman *et al.*, 2006
Retinol binding protein 4	*RBP4*	14*	TNB NBA	Rothschild *et al.*, 2000 Spotter *et al.*, 2009
Prolactin receptor	*PRLR*	16*	TNB NBA AP OR	Vicent *et al.,* 1998 Drogemuller *et al.*, 2001 van Rens and van der Lende, 2002 van Rens *et al.*, 2003
Leptin	*LEP*	18	Litter size	Chen *et al.,* 2004a

[1]Extracted from Bidanel (2011)
[2]SSC = *Sus Scrofa* Chromosome
[3]NBA = number piglets born alive; TNB = total number piglets born; TN = teat number; AP = age of puberty; NW = number weaned; LWW = litter weight at weaning; GL = gestation length; UC = uterine capacity; OR = ovulation rate.
*Candidate gene within reproductive QTL interval confidence.

Considering all candidate genes studies, it is evidenced that a there is a marked differences among them. These differences could be attributed to differences between resource populations, number of evaluated animals, mating systems, measured phenotypical traits and environmental influences (Buske *et al.*, 2006). Overall, this would tend to indicate that the causative mutation for reproduction traits is still unknown.

1.3.2.2. Candidate genes analysed in the MEIBMAP project

In the context of the MEIBMAP project, several candidate genes were analyzed for piglet survival traits and prolificacy related traits (Table 1.6).

Some of the candidate genes analyzed such as the *prolactin receptor* (*PRLR*) gene, the *dopamine β-hydroxylase* (*DBH*) gene and the *vascular cell adhesion molecule 1* (*VCAM1*) gene were associated with the newborn piglet vitality during the first hours after birth (Tomas *et al.*, 2006, Tomás *et al.*, 2006a,Ramirez *et al.*, 2008). Moreover, *PRLR* gene showed an effect on the ovulation rate (Tomas *et al.*, 2006). Nevertheless, these genes could not be associated with litter size traits. Unfortunately, no association could be detected for *parathyroid hormone-like hormone* (*PTHLH*) gene and the teat number although it was located on SSC5 where a QTL for teat number was identified (Martinez-Giner *et al.*, 2011).

Taula 1.6. Summary of candidate genes analyzed within the Meibmap project. The location in *Sus Scrofa* Chromosome (SSC), QTL associated with prolificacy related traits identified in the mapping position of the candidate gene, the polymorphisms changes at nucleotide and aminoacidic level (between brackets), the significant association with prolificacy traits and the reference of the manuscript published are indicated.

Candidate Gene	SSC	QTL[1]	SNP (aminoacid change)	Traits[2]	Reference
PRLR	16	-	C1217T(Leu406Pro) C1283A(Asp428Ala) G1439A(Lys480Arg) T1528A(Met510Leu) G1600A(Gly534Ser) G1789A(Gly597Ser)	CL** RT-1* HR-0† OS-1† TS†	Tomas *et al.*, 2006
DBH	1	-	T225C A463G(Thr155Ala) A510C T612C A616G(Lys206Glu) C744T	RT-1* TS† WB†	Tomás *et al.*, 2006a
VCAM1	4	-	T306A(Asn102Lys) C558T	TS* TU†	Ramirez *et al.*, 2008
PTPLH	5	TN	56C>T (Ser19Leu)	TN	Martinez-Giner *et al.*, 2011
BMPR1ƃ	8	-	G804C C852T C960T	NBA (1p)† NW (1p)†	Tomás *et al.*, 2006b
ESR1	1	-	PvuII polymorphism[3]	NBA* TNB*	Braglia *et al.*, 2006
MTNR1A	17	TNB/NBA	T162C	TNB* NBA*	Ramirez *et al.*, 2009
SLC9A3R1	12	TNB/NBA	A839G A259G	NBA* TNB*	Fernandez-Rodriguez *et al.*, 2010
NOS2	12	TNB/NBA	A662G A1791C C2192T	NBA* TNB*	Fernandez-Rodriguez *et al.*, 2010

[1]TN = Teat number; TNB = TNB = total number piglets born for 4 consecutive parities; NBA = number of piglets born alive for 4 consecutive parities.

[2]CL=number of corpus luteum; RT-1 = rectal temperature at 1h after birth; HR-0 = heart rate at birth; OS-1 = arterial oxygen saturation at 1h after birth; TS= time to the first suckle; WB = weight at birth; NBA (1p) = number of piglet born alive at first parity; NW (1p) = number of piglet weaned at first parity; TU = time to reach the udder; TN = sow teat number; NBA = number of piglets born alive for 4 consecutive parities; TNB = total number piglets born for 4 consecutive parities.

[3] Polymorphism described by Short et al. 1997.

***significant (p-value < 0.001); **significant (p-value < 0.01); *significant (p-value < 0.05); †suggestive (p-value < 0.1); ††no significant (p-value > 0.05).

Associations with the litter size traits could be determined for 5 of the candidate genes analyzed. Additive effects on prolificacy related traits were detected for the polymorphism PvuII (Rothschild *et al.*, 1996) on the *estrogen receptor 1* (*ESR1*) gene although it was discarded to be the causal mutation due to the effects associated with the Iberian and Meishan A alleles were significantly different (Braglia *et al.*, 2006). The *melatonin receptor 1A* (*MTNR1A*) gene, located within QTL for litter size traits on SSC17, was associated with litter size traits with additive and dominance effects that change depending on the season the year. Dominant effects for litter size traits were found to be during winter and autumn and they disappeared on summer and spring. However, the additive effects were determined on summer (Ramirez *et al.*, 2009). For *bone morphogenetic protein receptor type 1ƃ* (*BMPR1B*), only suggestively effects on litter size were determined and these could only be described during the first parity {{189 Tomás,A. 2006}}. Finally, the last two candidate genes analyzed the *solute carrier family* 9 (sodium/hydrogen exchanger) member 3 regulator 1 (*SLC9A3R1*) gene and the *inducible nitric oxide synthase 2* (*NOS2*) gene were significantly associated with the prolificacy related traits and, importantly, they were located within the confidence interval of the two epistatic QTL affecting litter size detected on SSC12 (Fernandez-Rodriguez *et al.*, 2010). The polymorphism detected within *SLC9A3R1* gene sequence was discarded to be the causal mutation while results for the *NOS2* gene were more conclusive. It has been suggested that *NOS2* haplotype could be the causal mutation underlying QTL2 on SSC12 although further investigations must be performed in order to extrapolate these results into commercial populations.

1.3.2.3. Candidate genes analyzed in this thesis

Four different candidate genes, three from the same family of genes, have been analyzed in the present thesis. All these genes have been selected due to its reported function related to reproductive processes (physiological candidate genes) and moreover, because they are mapped within the confidence interval of the QTL described in SSC13 (positional candidate genes) (Noguera *et al.*, 2009).

1.3.2.3.1. ITIH family: ITIH1, ITIH3 and ITIH4 genes

ITIHs belong to the superfamily of Inter-α-trypsin inhibitors (ITI-family) which is composed of a common light chain, designated as bikunin, and at least five closely related heavy chains (HCs) (ITIH-1, -2, -3, -4 and -5) (Himmelfarb *et al.*, 2004,Bost *et al.*, 1998,Salier *et al.*, 1996). Bikunin is bounded with the different heavy chains (ITIH-1, -2, -3) to form three protein complexes: the pre-α-inhibitor (PαI), the inter-α-inhibitor (IαI) and the inter-α-like inhibitor (ILαI) (Figure 1.3). It has been suggested that ITIH5 could also be bound with bikunin although there is no experimental evidence of this complex (Zhuo and Kimata, 2008,Zhuo and Kimata, 2008). On the other hand, ITIH4 could not be bounded with bikunin (Bost *et al.*, 1998,Salier *et al.*, 1996,Enghild *et al.*, 1989).

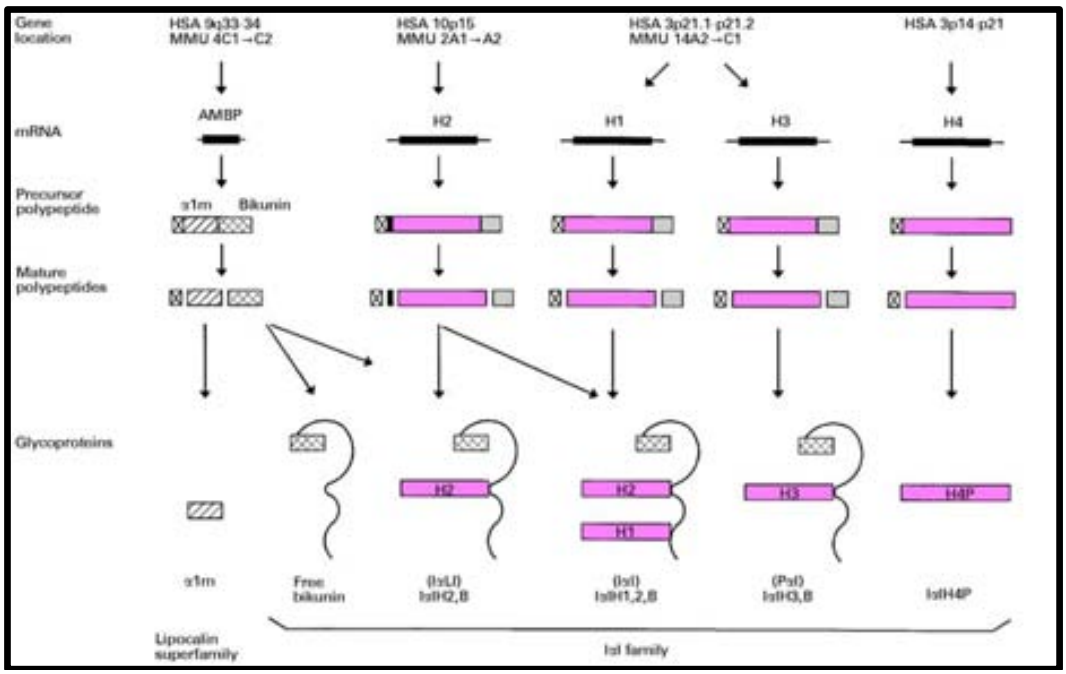


Figura.1.3. Schematic overview of ITI-family; from genes to protein complexes formation. (Extracted from (Salier *et al.*, 1996)).

The HCs are encoded by five genes which are located on two chromosomes. In pigs, ITIH-1, -3 and -4 form a gene cluster on SSC13 while ITIH-2 and -5 are located on another gene cluster on SSC10. ITIH-1, -2 and -3 are highly homologous at sequence level; for nucleotide as well as for aminoacidic sequences. On the other hand, ITIH4 is homologous with ITIH-1, -2 and -3 considering two-thirds of its sequence although the

rest is completely different. The high similarity among the HCs genomic sequences indicated that these genes could be formed as a consequence of gene duplication events (Zhuo and Kimata, 2008, Scarchilli et al., 2007)

The five HC precursor sequences have a von Willebrand-type (vWA) domain that functions as a binding site for integrins, collagen and proteoglycans, such as hyaluronic acid (Scarchilli *et al.*, 2007). This discovery was an important breakthrough that outlined the functional importance of the HCs. The HC-hyaluronan complexes are essential for ovulation and fertilization; specifically for the formation and the stabilization of the cumulus–oocyte complex (COC) during the preovulatory period (Chen *et al.*, 1990, Salustri *et al.*, 1989). Furthermore, ITIH-3 and -4 proteins have been proposed to be a class of acute phase proteins whose expression is activated as a consequence of an external stimulus such as an infection (Salier *et al.*, 1996, Gonzalez-Ramon *et al.*, 2000, Pineiro *et al.*, 1999).

ITIH-1, *-2*, *-3* and -4 genes are mainly expressed in liver and they are present in blood at considerably high concentrations (0.15 - 0.5 mg/ml) (Zhuo and Kimata, 2008). In pigs, these genes are also expressed in uterus. During embryo implantation, a 70-fold increase in bikunin expression has been reported (Hettinger *et al.*, 2001) whereas the *ITIH* expression remains constant throughout the oestrous cycle and early pregnancy (Geisert *et al.*, 2003). Thus, it is hypothesized that ITIHs may play an important role during pregnancy related to the stabilization of the glycocalyx at the endometrial–placental interface (Geisert *et al.*, 2003). Conversely, *ITIH5* has a different expression pattern compared with other members of the family. It is mainly expressed in placenta and its expression is minimal in liver (Himmelfarb *et al.*, 2004).

The porcine *ITIH-1*, *-3* and *-4* genes analyzed in the present thesis are large genes of about 14.03 Kb, 13.88 Kb and 11.48 Kb and they are composed by 21, 23 and 19 exons, respectively (Figure x). *ITIH-3* and *-4* encodes one transcript of 2,873 bp and 2,859 bp, respectively whereas *ITIH-1* can encode two transcripts of 2,595 bp and 2,873 bp. Finally, each *ITIH-3* and *-4* transcripts encode one protein of 888 and 731 aminoacids, respectively and the two *ITIH-1* transcripts two proteins of 865 and 902 aminoacids.

1.3.2.3.2. Mucin 4 (*MUC4*) gene

Mucins are large heterodimeric highly O-glycosylated glycoproteins commonly located on apical surfaces of many wet-surfaced epithelia of the body like the ocular surface, the respiratory, gastrointestinal and urogenital tracts, etc., being the principal component of the mucus. There are many mucins identified which are numbered based on its chronological order of discovery MUC 1, 2, 3, etc. These mucins can be divided into three subfamilies: the gel forming (secreted) mucins, the soluble mucins and the membrane-tethered mucins which include the studied one: the MUC4 (Moniaux *et al.*, 2001).

The porcine *MUC4* gene is located on SSC13, specifically on 13q41. It is a large gene with a total length of 31.12 Kb composed of 31 exons which are translated to a protein of 2123 aminoacids. Interestingly, porcine *MUC4* gene was located within the region associated with the Enterotoxigenic *Escherichia coli* (ETEC) F4ab/ac susceptibility and, moreover, it was proposed to be the most probable candidate gene (Jacobsen *et al.*, 2010). Two polymorphism described in porcine *MUC4* gene were associated with ETEC F4ab/ac susceptibility (Peng *et al.*, 2007,Jorgensen *et al.*, 2003). However, recent studies determined that porcine *MUC4* is distal from ETEC F4ab/ac susceptibility locus although the haplotype block between *MUC4* and ETEC F4ab/ac susceptibility is large (Rampoldi *et al.*, 2011).

MUC4, as all the mucins, presents anti-adhesive and anti-recognition properties (McNeer *et al.*, 1998a) which contributed to their main functions. They participate with the lubrication and with the protection of the mucosa from foreign substances and pathogens (Govindarajan and Gipson, 2010,Lagow *et al.*, 1999,Moniaux *et al.*, 1999). Therefore, *MUC4* gene is abundantly expressed in many ephitelia (Carraway *et al.*, 2009).

During pregnancy, *MUC4* is expressed on uterus of rodents and pigs although its expression during the peri-implantational period varies depending on the type of implantation in each species. In mice and rats, which have an invasive type of implantation, *MUC4* expression is downregulated to generate the receptive state for uterine implantation (McNeer *et al.*, 1998a,Carraway and Idris, 2001,Idris and

Carraway, 2000,Idris and Carraway, 1999,McNeer *et al.*, 1998b). Conversely, in pigs, where a non-invasive epitheliochorial placental attachment takes place, *MUC4* is upregulated in the uterus (Ferrell *et al.*, 2003). A protective role has been suggested for uterine *MUC4* which blocks the access of different substrates to the cell surface (Komatsu *et al.*, 1997). The endometrium is then protected from proteolytic activity of porcine conceptus (Ferrell *et al.*, 2003) and from microbial invasion (Carraway *et al.*, 2002) resulting in better uterine conditions for embryo development.

The importance of *MUC4* during reproduction processes is evidenced in many studies. In humans, it is associated with the development of endometriosis and endometriosis related infertility (Chang *et al.*, 2011). However, it has no effect with implantation failure (Koscinski *et al.*, 2006). On the other hand, MUC4 has been identified as a potential regulator of porcine placentation (Govindarajan and Gipson, 2010,Ostrup *et al.*, 2010).

1.4. TRANSCRIPTOME ANALYSES

1.4.1. Gene expression

The genetic information contained in the genome needs to be transcribed to generate the transcriptome and afterwards the transcriptome is translated to create the proteome. Both, transcriptome and proteome are the real effectors of the phenotypes. The transcriptome is defined as the total amount of transcripts that are expressed in a cell which is spatial-temporal and cell specific. In humans, the genome contains about 23,500 protein-coding genes which can encode more than 140,000 different transcripts, thus indicating that the transcriptome is more complex than the genome. In fact, the high phenotypic diversity observed across species cannot be explained only by the genomic variability observed at nucleotide sequence level. Therefore, it has raised the idea that this diversity must be generated by other mechanisms involving transcriptional and posttranscriptional modifications (Ruan *et al.*, 2004).

Deciphering the transcriptome profile of a cell or tissue helps in understanding the molecular basis of its biological function. An approach widely used in transcriptome analysis is to compare the mRNA expression level in specific tissues or cells under different conditions and/or treatments with the aim of identifying the differentially expressed (DE) genes with biologically relevant functions. Moreover, transcriptome analysis can be used to identify new biomarkers that could be used as diagnostic tools.

1.4.1.1. Gene expression analyses related to porcine reproduction

The pig genome has been estimated to encode 20,000–25,000 protein-coding genes (Groenen *et al.* 2011). Although the number of porcine transcripts has not been established yet, they are estimated to be as numerous as in humans.

Reproduction is a complex process that requires multiple physiological changes in different reproductive organs. These changes require a proper gene expression pattern which tissue and/or stage specific in order to assure an appropriate development of the reproduction function. Dysregulation of gene expression during reproductive processes leads to serious disorders that cause infertility such us endometriosis (Matsuzaki, 2011) and pre-eclampsia (Ishida *et al.*, 2011).

In pigs, as well as in other livestock species, the most studied reproductive process at transcriptome level is pregnancy. Specifically, the majority of works have been focused on the uterus where the gestation takes place. As a result, they have further increased our knowledge related to the molecular mechanisms that undergo during pregnancy though the research in pigs has just begun. During trophoblastic elongation and the early attachment to the uterine endometrium (Day 11-14 of the gestation), genes associated with cell motility, ATP utilization, cell growth, metabolism and intracellular transport have been reported as important for these processes (Ross *et al.*, 2009). Afterwards, at day 14 of the gestation, when the placentation process starts, up-regulation of genes involved in developmental process, transporter activity, calcium ion binding, apoptosis, cell motility, enzyme linked receptor protein signaling pathway, positive regulation of cell proliferation, ion homeostasis and hormone activity have

been reported in uterus. Conversely, genes related to oxidoreductase activity, lipid metabolic process and organic acid metabolic process are down-regulated (Ostrup *et al.*, 2010). Once the embryo is attached to the uterus, at day 30 of the gestation, genes related to enzymes involved in steroidogenesis and extracellular matrix remodeling and uterine secretory proteins have been suggested as important for the good development of the pregnancy (Ka *et al.*, 2008). Nevertheless, although the ovary is important during the gestation for the embryo survival and embryo development (Hunter *et al.*, 2004), little is known about the ovarian transcriptome. To our knowledge, only one study has undertaken to explore the ovarian transcriptome during pregnancy and it was by using the same resource IbxMe F_2 population as have used in the present thesis. Differentially expression regarding prolificacy levels was determined involving genes related to the immune system response activation and regulation against external stimulus, genes that regulate maternal homeostasis by complement and coagulation cascades and genes that are involved in lipid and fatty acid enzymes of metabolic processes (Fernandez-Rodriguez *et al.*, 2011).

On the other hand, some studies have focused on the transcriptome analysis in ovary during the oestrous cycle. Unfortunately, few works have attempted to describe the ovarian transcriptome during this process. In granullosa cells during ovarian follicles growth, the gene expression change showing a down-regulation of ribosomal protein, cell morphology and ion-binding genes and differential expression in genes related to lipid metabolism (Bonnet *et al.*, 2008). Ovarian expression of genes from *Insulin-like Growth Factor* (*IGF*) family as well as from *Epidermal Growth Factor* (*EGF*) family in are highly influenced by the oestrous cycle (Silva *et al.*, 2011). Interestingly, differential gene expression associated with follicogenesis was found between porcine lines selected for reproductive traits. In this sense, genes related to steroid biogenesis and tissue remodelling seems to have potential role during follicogenesis associated with reproductive traits (Caetano *et al.*, 2004,Gladney *et al.*, 2004).

Overall, porcine transcriptome analyses has evidenced that gene expression during reproductive processes are highly regulated. Thus, it is suggested that the mRNA expression levels could be used as potential biomarkers in porcine selection breeding programmes and/or as diagnostic tools. Nevertheless, further studies must be

performed in order to validate their use. Furthermore, the gene expression analyses could be used for the implementation of eQTL analyses which would allow identifying genome loci that regulate the expression levels. It would be of great interest in deciphering the reproductive traits but unfortunately, any eQTL associated with reproductive traits has been reported probably because of the high cost that these analyses involves.

1.4.2. Regulation of gene expression

As it is suggested in previous transcriptome analyses, the gene expression in cells/tissues is not constant. In fact, it needs precise temporal and spatial patterns. Thus, there are fundamental molecular mechanisms that regulate gene expression in order to assure a good respond of the cell to its environment and to the demands of the whole organism. The control of gene expression is operated by several mechanisms acting at different levels which can be divided into two steps: transcriptional control and post-transcriptional control. However, these steps are not independent; they are connected and coordinated (Dahan *et al.*, 2011,Orphanides and Reinberg, 2002). Overall, gene expression is controlled from the initiation of the transcription at nucleus to protein translation.

Traditionally, the gene expression regulation at transcriptional level has received the most attention because it is the most basic and important step of gene expression and it is easy to study with traditional established methods (Mata *et al.*, 2005). The importance of these regulation mechanisms is evidenced as the 5-10% of the total coding capacity of metazoans is dedicated to proteins that regulate transcription (Levine and Tjian, 2003). Mainly, gene expression at transcriptional level is regulated by proteins which can be divided into two groups: (1) sequence-specific DNA binding proteins (**transcription factors**) that mediated gene-specific selective transcriptional activation or repression and (2) proteins from large multi-protein RNA polymerase machines like TATA-binding proteins (Levine and Tjian, 2003). Furthermore, epigenetic mechanisms, which do not involve any change in DNA sequence, play an essential role

as regulators of transcription. Typically, epigenetic mechanisms comprise DNA methylation and histone modifications (Bell and Spector, 2011).

More recently, the importance of post-transcriptional regulation has emerged being essential in many biological processes and, moreover, a failure in post-transcriptional regulation can be the responsible for some diseases (Keene, 2007). The importance of these mechanisms has been evidenced by the fact that they provide a more rapid response to cellular signals and/or environmental stimulus than transcriptional regulation (Morris *et al.*, 2010). Although they are not well completely understood, it is known that RNA binding proteins (RBPs) and non coding RNAs (ncRNAs) are the main post-transcriptional mechanisms. Among ncRNAs, microRNAs (miRNAs) are clearly the most important post-transcriptional regulators of gene expression. Furthermore, other ncRNAs like piwi-protein-interacting RNAs (piRNAs), endogenous short interfering RNAs (endo-siRNAs) and long noncoding RNAs (lncRNAs) participate in the post-transcriptional control of gene expression (Strachan and Read, 2011). They regulate the gene expression through their binding to mRNAs sequences (Halbeisen *et al.*, 2008) influencing mainly the mRNAs stability and/or their translation (Huntzinger and Izaurralde, 2011).

1.4.2.1. miRNAs

MicroRNAs (miRNAs) are small ncRNAs of about 22 nucleotides that are abundant in animals, plants, fungi and even in virus. The first evidence of miRNA was reported by (Lee *et al.*, 1993) in *Caenorhabditis Elegans* by investigators that were studying phenomena related to RNA interference. However, it was not until the early 2000 that the first human miRNA was characterized (Pasquinelli *et al.*, 2000). Since then, the number of animal miRNAs described has been expanded substantially but the number of miRNAs described varies considerable between different species. For example, in humans, 1809 mature miRNAs have been reported whereas 618 mature miRNAs are known in pigs (Ensembl release 66 - February 2012).

miRNAs are key regulators of gene expression at post-transcriptional level inhibiting the translation process and/or modifying the stability of the mRNAs (Huntzinger and

Izaurralde, 2011). Each miRNA could regulate hundreds of different mRNAs and in mammals, it is estimated that miRNAs could regulated the 50% of all protein-coding genes (Krol *et al.*, 2010). Functional studies indicate that miRNAs participate in the regulation of almost every cellular process investigated including cell proliferation and differentiation, apoptosis and metabolism (Krol *et al.*, 2010, Bartel, 2009). Furthermore, the disruption of miRNAs expression pattern has been associated with different pathologies like cancer, neurodegenerative, inflammatory and cardiovascular diseases (Esteller, 2011).

It has been evidenced that miRNAs were present early on in the evolution of animals (Grimson *et al.*, 2008). During evolution, novel miRNA were acquired for particular lineages and, interestingly, the miRNAs acquired in a lineage is rarely lost in the descendent lineages (Heimberg *et al.*, 2010,Peterson *et al.*, 2009,Wheeler *et al.*, 2009,Heimberg *et al.*, 2008). This explains the high conservation in mammals' miRNAs. It seems that miRNAs have been important during the evolutionary process and particularly in the evolution of complexity of higher mammals (Bentwich *et al.*, 2005). Importantly, there is a direct correlation between the number of miRNAs and the complexity being the most complex species the ones who have more miRNAs in their repertoire (Berezikov, 2011). Nevertheless, there are miRNAs that are specie-specific which have been emerged in a particular specie. Novel miRNAs genes appear by different genomic sources. Processes like gene duplication, introns, pseudogenes, snoRNAs, tRNAs, transposable elements, antisense miRNA transcripts and de novo miRNAs transcripts from unstructured transcripts have been evidenced to participate in the formation of novel miRNAs (Berezikov, 2011). Once a miRNA is formed, it does not remain invariable at sequence level. There is sequence diversification implying mutation directly in the mature sequence regions and also mutations elsewhere in the precursor miRNA which lead to changes in hairpin structure. Mechanisms of miRNA sequence diversification are extensively reviewed by (Berezikov, 2011).

In animals, the genes encoding miRNAs are located in intergenic regions or in intronic regions being co-transcribed with the genes (Kim *et al.*, 2011). The biogenesis of miRNAs takes places through a multistep process that start in the nucleus with the transcription by RNA polymerase II of a precursor molecule called primary miRNA (pri-

miRNAs). Pri-miRNAs are folded in a hairpin structures which are cleaved by enzyme Drosha leading pre-miRNAs of about 70-bp which are transported to the cytoplasm by Exportin-5 protein. Afterwards, pre-miRNAs is processed by Dicer enzyme to obtain a ~ 20-bp dsRNA duplex. One strand of this duplex, representing a mature miRNA, is introduced to the miRNA-induced silencing complex (miRISC) which guides the miRNAs to mRNA complementary sequence to repress the gene expression (Figure 1.4). The strand that it is not incorporated in miRISC was initially thought that it was degraded but afterwards it was described that it remains on cells although at levels much lower than the strand that binds to miRISC. However, recently, it has been described a new molecular mechanism called arm switching (Okamura *et al.*, 2008) that define that the strand used to bind to miRISC can be switched between tissues and developmental time points (Chiang *et al.*, 2010,Ro *et al.*, 2007). In mammals, pre-miRNAs can also interact with the Argonaute 2 (AGO2) protein which can support Dicer enzyme and thus forming an additional processing intermediate called AGO2-cleaved precursor miRNA (ac-pre-miRNA) (Diederichs and Haber, 2007) (Figure 1.4).

miRNAs regulate their targets through a perfect Watson-Crick pairing between the seed region (from nucleotide 2 to 7 of the 5' end) of the mature miRNA and the target mRNA. Mainly, they downregulate their targets gene expression although recent studies have shown that they can also stimulate their target gene expression (Vasudevan, 2011). Several scenarios in which miRNAs contribute to gene regulation have been described (Bartel and Chen, 2004). miRNAs can act as genetic switches by fully downregulating their targets (Bagga *et al.*, 2005) or they can be neutral when the downregulation of mRNA targets is tolerated by cells or offset by feedback mechanisms. Nevertheless, it is widely accepted that most miRNAs act as micromanagers of gene expression by different mechanisms. They could fine-tune the gene expression level by adjusting the mean expression level of the target or they could act reducing the variance in the expression level of a target (**expression-buffering mechanism**). The importance of these mechanisms is providing system robustness to environmental perturbation (Wu *et al.*, 2009).

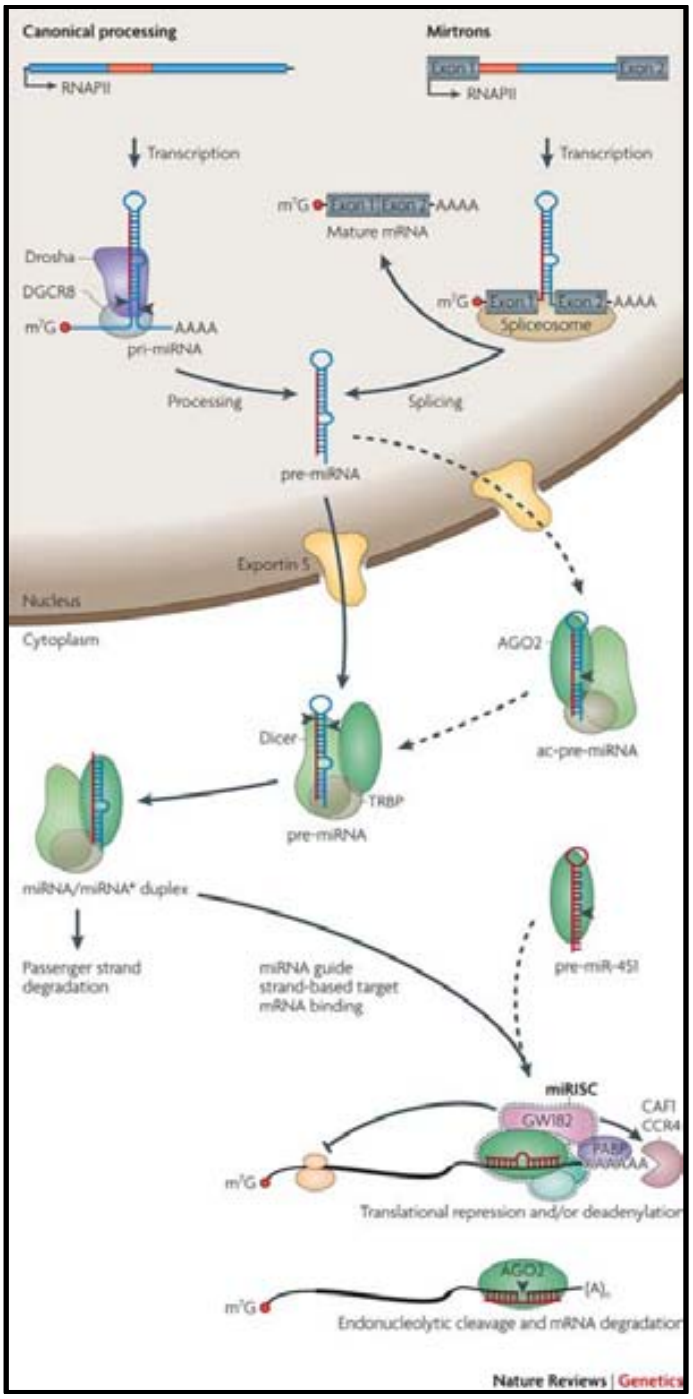


Figure 1.4. miRNAs biogenesis in animals. Extracted from (Knol *et al.*, 2002).

1.4.2.1.1. The role of miRNAs in reproduction

In mammals, post-transcriptional regulation of gene expression mediated by miRNAs is very important for female reproductive function. The relevance of miRNAs in reproduction is underlined by the fact that DICER, a key protein involved in the miRNA biogenesis, knockout mouse results in female sterility and the reproductive tissue

show multiple defects (Hong *et al.*, 2008,Nagaraja *et al.*, 2008). Thus, it was evidenced that DICER, and as a consequence miRNAs, are required for female fertility and reproductive tissue development. Furthermore, miRNAs are regulated by steroids hormones, which are essential for the regulation of the female reproduction (Cochrane *et al.*, 2011). As miRNAs are important for reproduction, dysregulation of miRNAs expression causes several pathologies. miRNAs has been associated with multiple type of cancers including ovarian, endometrial and cervical cancers (Hawkins *et al.*, 2011). The implication of miRNAs in other important reproduction diseases like premature ovarian failure (POF), endometriosis, uterine fibroids and preeclampsia has also been demonstrated (Zhou *et al.,* 2011, Hawkins *et al.*, 2011,Mouillet *et al.*, 2011,Ohlsson Teague *et al.*, 2009).

The importance of miRNAs in ovary has been reported in many studies (Li *et al.*, 2011,Baley and Li, 2012,Nothnick, 2012). The absence of DICER in murine ovaries has evidenced the key role of miRNAs in reproduction processes that take place in the ovary: folliculogenes, occyte maturation, ovulation and fertility (Hong *et al.*, 2008,Nagaraja *et al.*, 2008,Murchison *et al.*, 2007,Otsuka *et al.*, 2008,Lei *et al.*, 2010). However, the functions of ovarian miRNAs are started to be investigated mainly in human and mouse. Within the ovary, it is suggested that miRNAs expressed in somatic cells (i.e., granulosa cells which support the growth and development of the oocyte) highly influence on ovarian function (Baley and Li, 2012). By this way, miRNAs in granulosa cells control reproduction by regulating the expression of the major ovarian steroid hormones (progesterone, androgen and estrogen) (Sirotkin *et al.*, 2009). During folliculogenesis, in granulosa cells, Follicle Stimulating Hormone (FSH) stimulates miRNAs expression which may mediate changes in gene expression including the hormone production like progesterone (Yao *et al.*, 2010). Also in granulose cells but during the luteinization phase, miRNAs are regulated by luteinizing hormone (LH) where they may be important to regulate the gene transcription (Fiedler *et al.*, 2008) and to inhibit apoptosis (Carletti *et al.*, 2010). Moreover, it has been described that the lack of the miRNAs miR-17-5p and let-7b resulted in deficient corpus luteum and fertility suggesting a role for these miRNAs in the corpus luteum formation (Otsuka *et al.*, 2008).

Although the functions of miRNAs in the uterus is also a research area that is just starting to be explored, several works (mainly in humans and mice) have determined the importance of miRNAs in the majority of processes that are undertaken in the uterus. In particular, the role of uterine miRNAs has been focused on pregnancy when the uterus is of great importance. miRNAs can play important roles in the process of embryo implantation. Specifically, miRNAs participate in the regulation of genes associated with the inflammatory response in order to prepare the receptivity of the endometrium ((Pan and Chegini, 2008);(Chakrabarty *et al.*, 2007) for a successful implantation of the embryo. Moreover, miRNAs play roles in differentiation of uterine stromal cells in preparation for implantation by regulating the expression of genes involved in adhesion and extracellular matrix and by the cross-talk between the uterine epithelium and stroma (Hawkins *et al.*, 2011)(Pan *et al.*, 2007). In placenta, miRNAs are particularly abundant being important for the development of the placenta and for embryo survival (Cheloufi *et al.*, 2010). In sows, differential expression in placenta regarding the day of gestation has been determined suggesting the importance of miRNAs in pig placentation (Su *et al.*, 2010). Finally, it has been described that miRNAs are involved in parturition but further studies need to be performed in order to assess their functions (Kim *et al.*, 2011 ,Montenegro *et al.*, 2009).

1.4.3. Methodologies used to measure gene expression

The gene expression could be evaluated by a large variety of technologies. The measurement of gene expression was started to be individually for each gene. One of the first techniques used for this was Northern blot hybridization which can assess the expression level of transcripts by the relative intensity of the hybridization band. Afterwards, the reverse transcriptase quantitative PCR (RT-qPCR) technique appeared and this provided a more specific and accurate way to measure the expression levels of mRNAs. During the last decade, techniques like Expressed Sequence Tags (ESTs), Serial Analysis of Gene Expression (SAGE), Differential Display (DD), expression microarray and, more recently, high throughput sequencing, have been developed in order to

study the gene expression at whole-genome levels. Importantly, the most interesting results from expression data at whole-genome level needs to be further validated by each gene individually. To perform these validations, RT-qPCR has become the method most widely used. For simplification, only those techniques used in this thesis are described.

1.4.3.1. RT-qPCR

RT-qPCR is one of the techniques most used to measure expression mRNAs levels as well as other RNA species. During the procedure, the total amount of RNA needs to be converted to cDNA by a reverse transcription reaction. Then, the cDNA is amplified by PCR and in each amplified cycle, the product is quantified by the use of a fluorophore. For this, the fluorescence emitted at each cycle is registered being proportional to the amount of the amplified product. By this way, the PCR reaction can be plotted showing the typical pattern as it is shown in Figure 1.5. The quantification of the expression level for each sample is performed in the exponential phase by using the crossing threshold (Ct) (Figure 1.5). The Ct is the number of cycles that it takes each PCR reaction to reach an arbitrary amount of fluorescence called threshold. Thus, the Ct is inverse proportional to the expression level. In this way, if the expression level is low, the Ct will be high.

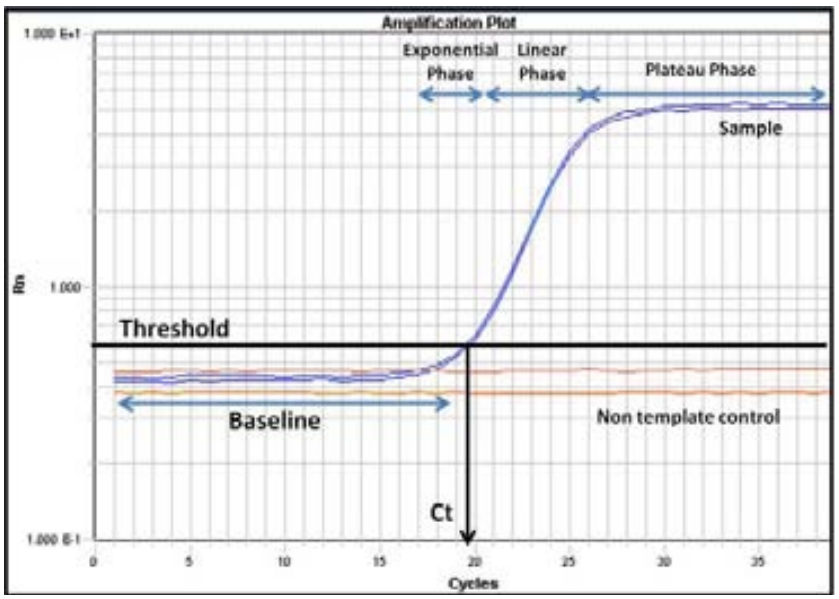


Figure 1.5. RT-qPCR amplification plot of one sample and the non template control.

The fluorescence needed for the quantification of the gene expression is based on different technologies including (i) probe sequences that fluoresce upon hydrolysis (TaqMan®; Life Technologies) or hybridization (HybProbes; Roche); (ii) fluorescent hairpins hybridization (LUX; Invitrogen and Molecular Beacons; Sigma-Aldrich, IDT); or (iii) intercalating dyes (SYBR Green). In the present thesis, the SYBR Green is the technology used. It is an intercalating dye that emits fluorescence when it binds to double-stranded DNA. Following primer-mediated replication of the transcript during PCR, multiple molecules of SYBR Green are bound to the PCR product and emit a strong fluorescent signal.

Analysis of RT-qPCR data can be performed by measuring the absolute expression level (i.e. number of copies of a specific gene transcript per sample) or by the relative expression level (i.e. the transcript of a specific gene is expressed twice in sample 1 than in sample 2). So far, the relative quantification is the method most used in expression studies. Within the relative quantification, the $2^{-\Delta\Delta CT}$ Ct method (Livak and Schmittgen, 2001) is the most used. This method is based under the assumption that the efficiency of the RT-qPCR is of 100% and the need of using a reference gene which expression is stable between the analyzed samples to correct for any difference in sample loading. Normalized expressions are then made relative to a calibrator sample which is randomly chosen.

1.4.3.1.1. Adaptation of RT-qPCR for analysis of miRNAs

The RT-qPCR for measure miRNAs expression follows the same concepts as explained above for determining mRNA expression. However, the main difference falls on the short length of miRNAs. They are only 22 nucleotides long which is the same size as conventional RT-qPCR primers. Thus, it is challenging to design RT-qPCR primers for miRNAs. Several methodologies have been developed in order to overcome this issue which are summarized in Table 1.6.

Table 1.6. Summary of the RT-qPCR methods described for measure miRNA expression.

	Stem-loop RT-qPCR[1]	poly(A) miRNA-based RT-qPCR[2]	Universal RT microRNA PCR[3]	miR-ID[4]
miRNA preparation		miRNA polyadenilation	miRNA polyadenilation	miRNA circularization
Primer RT reaction	Specially designed loop-primer	Poly(T) primer	Tagged poly(T) primer	Specific circular miRNA primer
Primers RT-qPCR reaction	miRNA specific primer + universal primer	miRNA specific primer + universal primer	two specific primers spiked with LNA®	miRNA specific 5'-overlapping primers
PCR detector	Taqman Probes	SYBR Green	SYBR Green	SYBR Green
Control reaction specificity	No	Yes	Yes	Yes

[1] (Chen *et al.*, 2005); [2] (Shi and Chiang, 2005); [3] Busk, 2010; [4] (Kumar *et al.*, 2011).

These RT-qPCR miRNAs methodologies have been widely used in miRNA expression studies. The poly(A) miRNA-based RT-qPCR (Shi and Chiang, 2005) and the Universal RT microRNA PCR (Busk, 2010) (Exiqon, Denmark) are the most suitable methods when the amount of sample is limiting and when the miRNA concentration is low because they only require a single reverse transcription. Although stem-loop RT-qPCR (Chen *et al.*, 2005,Mestdagh *et al.*, 2008) (Life Science) and miR-ID (Kumar *et al.*, 2011) methods allow can be scaled in order to perform multiplex studies, the number of miRNAs which can be studied is limited by the RT primers used; 384 and 20 miRNAs, respectively.

There are some issues that have to be considered related to RT-qPCR methods used for measuring miRNA expression. In case of stem-loop RT-qPCR method (Chen *et al.*, 2005) (Life Science), the main inconvenient is that specificity of the RT-qPCR reaction cannot be controlled. Although a TaqMan probe (Life Science) is used, it does not assure the specificity of the reaction due to the TaqMan® probe binds to a cDNA sequence which is synthesized during the reverse transcription reaction. By this way, the RT primer could bind to a different sequence, which cannot be the miRNA of interest. Then, the TaqMan® probe will bind to an unspecific amplicon and it will be

indistinguishable from the desired PCR product. The poly(A) miRNA-based RT-qPCR (Shi and Chiang, 2005) only uses one specific miRNA primer to perform the RT-qPCR reaction. The design of miRNA specific primer is limited and it would be a problem when the discrimination between closely related miRNAs differing only in one nucleotide is needed. Finally, Universal RT microRNA PCR (Busk 2010) (Exiqon, Denmark) is very specificity by using two miRNA specific primers spiked by LNA. However, they usually show low RT-qPCR efficiencies which can lead inaccurate quantification (Balcells *et al.,* 2011).

1.4.3.2. Microarrays

Expression microarrays have provided much important information about how transcripts are expressed in different cell types and tissues under specific conditions at genome wide level. An expression microarray consists of hundreds of probes which are immobilized on a solid substrate, chip or array. The probes used are complementary to the transcripts that are under investigation. There are many commercial available microarrays which have been designed for specific species including pig. In these microarrays, the number of transcripts to be evaluated is predetermined but, as an alternative, it is possible to use custom design arrays in order to only insert on the array the interested transcripts.

Among the commercial available microarrays, the Affymetrix 24K Genechip® Porcine Genome Array (Affymetrix, Santa Clara, CA, USA) was chosen in the present thesis. Total RNA is labelled and afterwards is hybridised on the array being the amount of signal emitted proportional to the expression level. Then, the labelled RNA molecules are hybridized to the arrays and they are scanned to determine, for each probe, the emitted fluorescence which represents the expression level of each miRNA.

Once the fluorescent intensity for each probe is registered, a quality control needs to be applied to assess if the process has performed properly. For this, signal intensities obtained are evaluated in order to detect if the hybridisation to the array has done uniformly or there are array zones in which there is no hybridization or it is excessively.

The quality control in commercial arrays is performed by comparing the signal intensities of endogenous gene probes which have been inserted into the array. Thus, if the signals of quality control probes are similar among the arrays analyzed, then it is possible to compare the different arrays. However, there are many variation sources during the microarray procedure that do not represent true biological variation and they must be removed. For this, a normalization method is needed. Several normalization methods have been developed during the last 10 years (i.e. scale normalization, quantile normalization, Lowess normalization, etc). The selection of the best normalization method is taking into consideration the data obtained. It very important to select an adequate normalization method due to it has a great impact on the results and, therefore, in the biological interpretation. After data is normalized, statistical method can be applied. Traditionally, mRNA microarrays are performed to determine differential mRNA expression under specific conditions. For this, the most used statistical tests are t-test or Analysis of Variance (ANOVA) for each probe. Nevertheless, as multiple tests are performed at the same time, statistical correction like False Discovery Rate (FDR) must be applied to eliminate the false positive results. Finally, genes that have been determined as differential expressed (DE) are used to perform an *in silico* functional analysis in order to identify which biological processes are enriched and to identify the biological pathways where the DE genes are involved.

Several softwares have been designed with the purpose of analyzing microarray data. Among these, one of the most used is Bioconductor (Gentleman *et al.*, 2004) which is implemented in R (http://cran.r-project.org/). By using this software, a complete microarray analysis can be done; from quality control to differential expression tests. On the other hand, *in silico* functional analyses are performed by using bioinformatics tools that allow a functional annotation of the genes through Gene Ontology (GO) analysis by using softwares such as Database for Annotation, Visualization and Integrated Discovery (DAVID), WEB-based GEne SeT AnaLysis Toolkit (WebGestalt), etc and the integration of these genes into genetic networks by using *Ingenuity pathway software* (IPA), FunCluster, FunNet, etc. *In silico* functional analyses are complicated when it is worked with livestock species like pig due to there is a lack of gene annotation and most of the pathways where genes are involved are still unknown.

Thus, to perform accurate functional analyses, the human annotation is used because genes are better annotated are the knowledge of the genetic networks is higher. However, these analyses are performed under the assumption that genes are highly conserved among mammals.

1.4.3.2.1. Differences between mRNAs and miRNAs microarrays

miRNA microarray follows the same bases as described above for mRNA microarrays. There are available several commercial miRNA microarrays which are based on described miRNAs, mostly focussed on human, mouse and rat. Do to the fact that miRNAs are highly conserved across mammals; these commercial miRNAs microarrays could be used in a wide range of species including the pig although specie-specific miRNAs obviously cannot be evaluated.

The main difference between miRNAs and mRNA is the small size of miRNAs (~ 22 nucleotides). Thus, the probe design in miRNA microarrays could not be performed as for mRNA microarrays. In fact, miRNA probe design has become a challenge due to the fact that the sequence for designing probes is very short. Moreover, the melting temperature (Tm) of different probes must be normalized in order to achieve optimal hybridization during the assay at one temperature. Different technologies have been developed in order to achieve these issues (Li and Ruan, 2009, Yin *et al.,* 2008). One of the most used, and used in the present thesis, has been developed by the company Exiqon (Denmark). They have designed LNA™ probes by mixing nucleotides with Locked Nucleic Acid (LNA™) nucleosides in which the ribose ring is “locked” by a methylene bridge connecting the 2’-O atom and the 4’-C atom (Exiqon). By this way, the probes became highly specific providing the distinction between closely related microRNA family members that differ only in one nucleotide. Also highly sensitivity is achieved by using these LNA™ probes that is suitable for the detection of miRNAs highly and lowly expressed at the same time in the assay. In addition, by using LNA™ probes, the Tm could be normalized and increased by adjusting the LNA contents and lengths of the probes.

1.4.3.3. High-throughput sequencing (RNA sequencing)

During the last 5 years, the development of new tools for high throughput sequencing (HTS) of DNA molecules has revolutionized the molecular genetics. These new sequencing methods provide thousands of sequences in a very short time. In addition to genome sequencing applications, HTS can be used as a new tool for transcriptome analysis which is known as RNA sequencing (RNA-seq).

Up to now, HTS can be performed by using different technologies that can be divided into two groups according the sequencing methodology used (Strachan and Read, 2011): (1) Sequencing by synthesis of PCR-amplified DNA (massively parallel pyrosequencing (454 pyrosequencing (Roche/454), Reversible dye-terminators sequencing (Illumina/Solexa) and Ion Semiconductor Sequencing (Life Science/Ion Torrent)) (2) Sequencing by ligation of PCR amplified DNA (Sequencing by Oligonuleotide Ligation Detection (SOLiD) (Life Technology)) and (3) Single molecule sequencing (Single Molecule Real Time (SMRT®)(Pacific Bioscience)). Currently, Nanopore Company is developing a new sequencing technology called Hybridization-Assisted Nanopore Sequencing (HANS) which is based on single molecules sequencing. Importantly, the development of single molecules sequencing methods are of great interest due to they can reduce the cost of the whole-genome DNA sequencing by more than four orders of magnitude.

Massively parallel 454 pyrosequencing was the first HTS technology developed by 454 Life Science company but it is now commercialized by Roche. In the present thesis, it is the technology used to perform the miRNA transcriptome profile. The workflow starts, as in the others HTS technologies, with the preparation of a small RNA library. The small RNAs (15 - 30 nucleotides) are isolated from the total RNA (i.e., by using denaturing polyacrylamide gel). Afterwards, small RNAs are ligated with a 3' and a 5' linker in order to perform the reverse transcription reaction where the cDNA is obtained. This cDNA is amplified by a PCR reaction to increase the number of small RNAs molecules and, importantly and specific for the HTS method chosen, to integrate the 454-specific sequencing adaptors to PCR products which are necessary to perform the sequencing protocol. Then, the 454 pyrosequencing (454/Roche) protocol starts

(Figure 1.6). The first step is denaturing the cDNA and bound the cDNA into beads through the specific adaptors of the DNA which are recognized by the beads. The cDNA-beads binding is performing under conditions that favour one fragment per bead. Thereafter, the beads with the captured DNA are clonally amplified by an emulsion PCR (emPCR). For this, beads are placed in a water-in-oil emulsion, forming hundreds of thousands of PCR microreactors. The result is 10 million copies of a unique DNA template in each bead. When the emPCR is finished, the beads are deposited into individual wells of a fiber-optic slide (PicoTiterPlate®, Roche/454) which contains 1.6 millions wells that contain the reagents for the sequencing reaction which is specifically a pyrosequencing reaction (Ronaghi *et al.*, 1998). Briefly, in pyrosequencing, a DNA polymerase synthesizes a DNA chain by using a single-stranded DNA template. Instead of being all 4 dNTPs, individual dNTPs are provided sequentially. When the dNTP provided is incorrect, it is degraded by the enzyme apyrase. However, when the correct dNTP is provided, it is incorporated into the DNA chain producing a pyrophosphate (PPi) which is used by the enzyme ATP sulfurylase to generate ATP which, in turn drives a luciferase reaction to produce light. The light emitted is captured to compose the sequence.

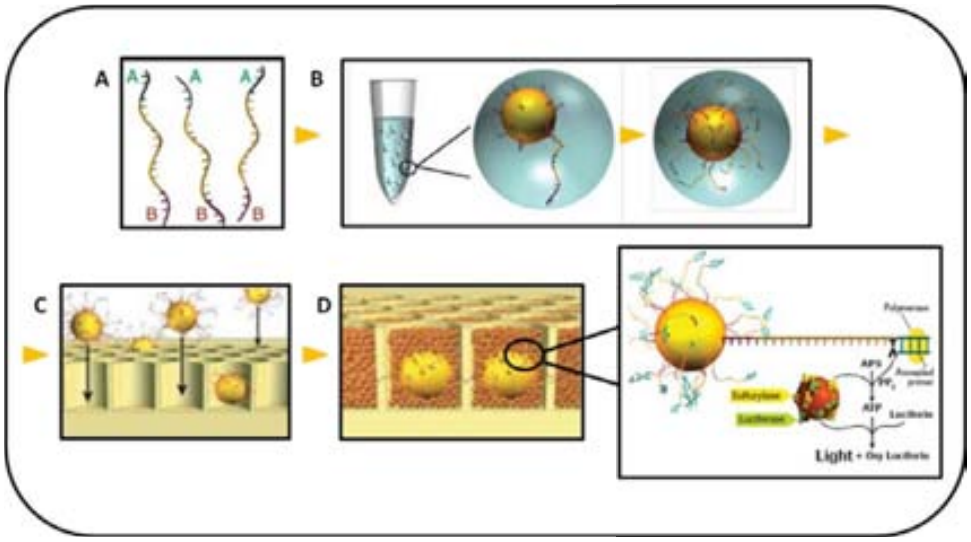


Figure 1.6. miRNAs preparation for massively parallel 454 pyrosequencing (454/Roche). (A) miRNA cDNAs ligated with 454 specific adaptors. (B) miRNA cDNAs are bound into beads under conditions that favour one sequence per bead. Thereafter, the beads with the captured cDNA are clonally amplified by an emulsion PCR in a water-in-oil emulsion obtaining million copies for each fragment. (C) Diposition of the beads deposited into individual wells of a fiber-optic slide (PicoTiterPlate®, Roche/454). (D) Pyrosequencing reaction.

After performing a HTS methodology, thousands of sequences are obtained. Then, a bioinformatic analysis is performed to identify miRNAs sequences. The sequences for each miRNA are counted being the measure of the expression level. During library preparation, many factors like RNA ligase preferences, reverse transcription reaction and PCR based amplification could bias the HTS experiment (Linsen *et al.*, 2009). Thus, a normalization method is necessary to remove this library preparation bias. However, although a satisfactory correction model has not yet been identified, scaling to library size is the standard procedure to normalize small RNA-seq experiments (Marques *et al.*, 2010).

Importantly, compared with microarray experiments, HTS methods provide direct access to the sequence. By this way, they allow the identification of de novo small RNA transcripts. Moreover, different events can also be analyzed. For example, the length variability in miRNAs (isomiRs), RNA editing events and polymorphism within small RNAs can also be identified. Moreover, the miRNA expression profile could be determined in species whose genome has not been sequenced. Nevertheless, it requires a considerable amount of time for data generation and data analysis.

2. OBJECTIVES

The present research has emerged from the MEIBMAP project (AGL2000-1229-C03 and AGL2004-08368-C03), a coordinated project between IRTA, INIA and UAB whose main goal was the characterisation of the genetic architecture of economically important traits affecting reproduction in pigs. An experimental F_2 intercross between Iberian boars and Meishan sows was generated and QTL affecting several reproductive traits were identified (Noguera *et al.*, 2009, Fernandez-Rodriguez *et al.*, 2010, Rodriguez *et al.*, 2005). In particular, two QTL for the number of piglets born alive (NBA) and the total number of piglets born (TNB) were located on porcine chromosomes 13 (SSC13) and SSC17 (Noguera *et al.*, 2009). Moreover, for these traits, several epistatic QTL were reported evidencing that prolificacy related traits are the result of a complex network of interacting genes. Based on the previous knowledge that evidenced the effect of several genomic regions on litter size, the general objective of the present thesis was to increase the understanding of the genetic basis underlying the prolificacy QTL in the IbxMe F_2 population. This thesis has been performed in the frame of the project MIRPIGVIR (AGL2007-66371-C02 and AGL2010-22358-C02) founded by the Ministry of Science and Innovation.

The specific objectives of the current thesis are:

1) Characterization of four candidate genes for prolificacy QTL on SSC13 described by (Noguera *et al.*, 2009): the *inter-alpha-trypsin inhibitor heavy chain* (*ITIH*) *-1*, *-3* and *-4* and the *mucin 4* (*MUC4*).

2) Identification of thee genes differentially expressed in uterus between high and low prolificacy IbxMe F_2 sows at 30-32 days of pregnancy.

3) Characterization of the uterine and ovarian microRNAome at 30-32 days of gestation in IbxMe F_2 sows regarding prolificacy levels.

4) Development of a new RT-qPCR methodology for measuring miRNA expression levels.

3. ARTICLES AND STUDIES

3.1. CANDIDATE GENES ANALYSES ON SSC13

3.1.1. Characterisation of *ITIH* gene cluster for prolificacy QTL on SSC13

Sequencing and gene expression of the porcine *ITIH* SSC13 cluster and its effect on litter size in an Iberian x Meishan F_2 population

Balcells, I., Castelló, A., Noguera, J.L., Fernández-Rodríguez, A., Sánchez, A. and Tomás, A.

Animal Reproduction Science (Accepted)

Sequencing and gene expression of the porcine *ITIH* SSC13 cluster and its effect on litter size in an Iberian x Meishan F_2 population

I. Balcells[1,2,*], A. Castelló[1,2], J.L Noguera[3], A. Fernández-Rodríguez[4], A. Sánchez[1,2] and A. Tomás[5]

[1]Departament de Ciència Animal i dels Aliments, UAB, 08193 Bellaterra, Spain

[2]Centre de Recerca Agrigenomica (CRAG), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[3]Genètica i Millora Animal, IRTA-Lleida, 25198 Lleida, Spain

[4]Departamento de Mejora Genética Animal, SGIT-INIA, 28040 Madrid, Spain

[5]Program Infection and Immunity, FISIB, 07110 Mallorca, Spain

[*]Correspondence author

ABSTRACT

The aim of the present study was to identify polymorphisms and to analyze endometrial gene expression of the porcine SSC13 *ITIH* cluster that could explain differences in prolificacy of 255 F_2 sows derived from an Iberian (Ib) x Meishan (Me) intercross in which QTL for the number of piglets born alive (NBA) and total number of piglets born (TNB) were previously detected on this chromosome. Sequencing of *ITIH-1, -3,* and *-4* mRNAs was done and several polymorphisms segregating within the IbxMe population were found in all three genes. Significant associations with NBA were found for two SNPs from *ITIH-1*, four from *ITIH-3*, and four SNPs from *ITIH-4* (p<0.05). Haplotypes for the significant SNPs were calculated by segregation analysis and a marker assisted association test indicated that the alleles coming from the Meishan breed had a favorable effect on NBA for all three genes. Interestingly, some of the significant SNPs were located within the von Willebrand domain of the ITIH proteins, the binding site of molecules essential for the synthesis of the extracellular matrix during cumulus expansion. Gene expression analyses also revealed differences in the expression level of the *ITIH-3* gene regarding the prolificacy performance (high or low) and the uterus sample (apical or basal).

KEYWORDS

ITIH-1, ITIH-3, ITIH-4, litter size, polymorphism, expression

INTRODUCTION

Reproductive traits, such as prolificacy, are complex traits of economical interest for swine breeders. Most of them, however, display low heritabilities being difficult to improve them through conventional selection programmes. In the last decades, a big effort has been made by the scientific community to identify and characterize the genes and/or genomic regions involved in reproductive processes through quantitative trait loci (QTL) approaches (Spötter and Distl, 2006) which could potentially be used as genetic markers in marker assisted selection (MAS) programmes. In a previous study made by our group, we reported the most significant QTL for prolificacy at the genome-wide level (Noguera et al, 2009). Two QTL for the number of piglets born alive (NBA) and total number of piglets born (TNB) were identified on chromosomes 13 (SSC13) and SSC17, as well as multiple epistatic interactions throughout the porcine genome in an Iberian x Meishan (IbxMe) F_2 population (Noguera et al. 2009). The success of the Noguera et al. (2009) study is probably due to the use of two highly divergent European and Asian breeds which present extreme phenotypes for prolificacy traits. The Meishan is considered one of the most prolific breeds (average of 14.3 piglets born alive per parity, Bidanel et al. 1989) whereas Iberian is considered a very low prolificacy breed (average of 7.0 piglets per parity, Silió et al. 2001). Additionally, these breeds are also highly divergent at the genetic level (Alves et al. 2003). By examining the genes located within the confidence intervals of the QTL described above, we have identified the *inter-α-trypsin inhibitor heavy chains* (*ITIH*) *-1, -3,* and *-4* cluster as a potential candidate genomic region to explain the NBA and TNB QTL on SSC13 (Baskin et al. 1998).

ITIHs belong to the superfamily of kunitz-type protease inhibitors and are assembled from five distinct polypeptides that include four heavy chain precursor proteins (ITIH-1, -2, -3, and -4) and a light chain, designated as bikunin (Salier et al. 1996; Bost et al. 1998). The four heavy chain precursor sequences have a von Willebrand–type (vWA) domain that functions as a binding site for integrins, collagen and proteoglycans, such as hyaluronic acid (Scarchilli et al. 2007). During the pre-ovulatory period of cumulus expansion, the cumulus-oocyte complex (COC) volume is greatly increased by secretion of large amounts of hyaluronic acid to the extracellular matrix (Salustri et al. 1989;

Chen et al. 1990). In pigs, the placental form of implantation is non-invasive. Protease inhibitors are secreted by the endometrium to prevent an excessive invasion of trophoblast into the thick glycocalyx on the apical uterine surface epithelium (Roberts et al. 1993; Stroband and Van der Lende, 1990).

In pigs, embryo implantation takes place on days 15–18 of pregnancy. During this period, a 70-fold increase in *bikunin* expression has been reported (Hettinger et al. 2001) whereas the *ITIH* expression remains constant throughout the oestrous cycle and early pregnancy (Geisert et al. 2003) which indicates that bikunin and the ITIHs may play an important role during the period of conceptus attachment to the uterine luminal surface (Bhanumathy et al. 2002).

The aim of the present study was to identify polymorphisms and to analyze endometrial gene expression of porcine *ITIH-1, -3* and *-4* genes that could explain differences in prolificacy of sows derived from an F_2 IbxMe intercross in which QTL for NBA and TNB were detected at positions 50 and 55 cM on SSC13, respectively (Noguera et al. 2009).

MATERIAL AND METHODS

Animal material and phenotypic measurements

An F_2 population was generated by crossing 3 Iberian males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) to 18 Meishan females (Domaine du Magneraud, INRA, France). Eight boars and 97 sows from the F_1 generation were mated to obtain the F_2 progeny in the experimental farm Nova Genètica S.A (Lleida, Spain). Among F_2 sows, 255 F_2 sows were randomly selected for mating and generated a total of 881 parities with an average of 3.45 parities per F_2 sow. At farrowing, the total number of piglets born (TNB) and the number of piglets born alive (NBA) was determined. An average of phenotypic data for TNB and NBA were 8.69 (±3.04) and 9.02 (±3.10), respectively. In the fifth reproductive cycle, sows were slaughtered at 30-32 days of gestation and the number of embryos (NE) was recorded. At slaughter, 2 uterus

samples were collected: apical uterus (proximal to the ovary) and basal uterus (close to the cervix).

Sequencing of *ITIH-1*, *-3* and *-4* mRNAs

Due to the fact that no fresh tissue samples from the parental population were available for cDNA amplification and polymorphism identification, we used uterus samples from different porcine breeds that were available in our laboratory for this purpose. Therefore, total RNA was extracted from 11 sows (3 Meishan, 2 Iberian, 2 Vietnamese, 2 Landrance, 1 Large White and 1 Piétrain) by using the RiboPure™ kit (Ambion). Synthesis of cDNA was performed with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) from 1µg of total RNA in 20µl reaction. Primers for *ITIH-1*, *-3* and *-4* genes were designed based on pig sequences available at the pig genome sequence (*Sus scrofa*, Ensembl release 61 - Feb 2011, accession numbers ENSSSCT00000012530, ENSSSCT00000012532 and ENSSSCT00000012534, respectively). Five, six and four different primer pairs were used for the amplification of the coding region of *ITIH-1 (PCR1-5)*, *ITIH-3 (PCR1-6)* and *ITIH-4 (PCR1-4)* genes, respectively (Supplementary Table 1). PCR reactions were performed in a 15 µl final reaction mixture containing 1.5-2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1U of DNA polymerase, 0.2 µM of each primer, and 1-2 µl of cDNA (Supplementary Table 1). Thermal profiles for each PCR consisted of 35 cycles of 95ºC for 1 min, Tm for 1 min and 72ºC for 1 min. Specific Tm were used for each primer pair (Table 1). PCR products were enzymatically purified using 0.4 units of *Exonucease I* (Quimigen) and 0.2 units of *Alkaline Phosphatase* (Quimigen) during 1 hour at 37ºC followed by inactivation at 85ºC for 15 min. Purified PCR products were sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction kit v3.1 (Applied Biosystems). Sequencing reactions were filtered using Montage SEQ96 Sequencing Reaction Cleanup Kit (Millipore Corporation) and electrophoresed in an ABI PRISM 3730 Genetic Analyzer (Applied Biosystems). Polymorphisms were detected by aligning the resulting sequences with the SeqScape® v2.5 software (Applied Biosystems).

In order to identify which of the detected polymorphisms were segregating within the Iberian and Meishan parental populations, several primers were designed based on the available porcine sequence to amplify, from genomic DNA, the different polymorphic fragments of *ITIH-1* (PCR6-8), *ITIH-3* (PCR7-9) and *ITIH-4* (PCR5-6). PCR reactions, thermal profiles, PCR purification and sequencing were performed as abovely explained.

Typing of polymorphisms

Sixteen polymorphisms (three in *ITIH-1*, nine in *ITIH-3* and four in *ITIH-4;* Table 2*)* were genotyped in the IbxMe population to determine the effects of these polymorphisms in prolificacy traits.

A pyrosequencing (Ronaghi et al. 1998) protocol was carried out in multiplex when possible (for *ITIH-3* SNPs 1332, 1341, 1362; and for *ITIH-4* SNPs 826, 840, 843 and 1027). For those SNPs for which a pyrosequencing protocol could not be set up, we used a primer extension analysis strategy based on the SNapSHOT Multiplex (Applied Biosystems) kit (for *ITHI-1* SNPs 744, 2164, 2623; and for *ITIH-3* SNPs 886, 888, 909, 1736, 1867, and 1872). The nucleotide sequences of primers and probes used for both methodologies as well as specific reaction conditions are described in Supplementary Table 1. Pyrosequencing and primer extension reactions were analysed in a PSQ-H96 pyrosequencer (Pyrosequencing AB, Uppsala, Sweden) and in an ABI PRISM 3730 DNA analyzer (Applied Biosystems); respectively.

Association analysis

The linkage map of porcine chromosome 13 (SSC13) was constructed using the *build* option of the CRIMAP 2.4 software considering the haplotypes of significant SNPs of the *ITIH-1, -3* and *-4* genes (Green et al. 1990). The effects of the genotypes and haplotypes in prolificacy traits were tested by a univariate mixed-inheritance animal model. Analyses were performed with the Qxpack v4.1 software (Pérez-Enciso and

Misztal, 2004); the *snp* option was used for each individual SNP and the *ld_fix* option for haplotypes.

$$y = H_i + O_j + u_k + p_k + \Sigma_k \lambda_{ik}\, a_k + e_{ijkl} \text{ (model 1)}$$

Association model (model 1) was used to test the individual allelic effects of each SNP on prolificacy traits (Kadarmideen et al. 2011). y_{ijk} was the ijkth observation for NBA and TNB, H_i was the ith year-season fixed effect, O_j was the jth parturition number fixed effect, u_k was the random polygenic effect of the kth individual, p_k was the environmental permanent effect for the kth individual, a_k was the additive effect of the SNP, λ_{ik} is a variable that indicates the number of copies (0, 1, or 2) of the k^{th} allele (note that k = 1 or 2 in the analysis of SNPs) and e_{ijkl} is the random residual term. Dominance effect was not included in model 1 because it was not significant. The effect size of each significant SNPs of the *ITIH* SSC13 cluster was calculated as Pryce et al. (2011).

To test whether the effect observed for *ITIH-1*, *-3* and *-4* SNPs were not due to linkage disequilibrium with the NBA and TNB QTLs found on SSC13 by Noguera et al (2009) in the same population, a marker-assisted association test (MAAT, Zhao et al. 2003) was performed (model 2). This model allows differentiating if the observed effect of a gene is due to linkage disequilibrium with a nearby QTL or it is independent. To gain information, haplotypes of the significant SNPs were determined by segregation analysis within the IbxMe population and included in the model 2. Therefore, model 2 was as follows:

$$y = H_i + O_j + u_k + p_k + c_a\, a_1 + \Sigma_k \lambda_{ik}\, b_k + e_{ijkl} \text{ (model 2)}$$

where a_1 is the QTL additive effect, C_a is the probability of the individual to be homozygous for Iberian origin alleles minus the probability of being homozygous for Meishan alleles at each QTL position (every 1 cM), b_k is the additive effect of the haplotype, λ_{ik} is a variable that indicates the number of copies (0, 1, or 2) of the k^{th} allele (note that k = 1 or 2 for *ITIH-1* and *ITIH-4* haplotypes and k = 1, 2, or 3 when *ITIH-3* haplotype is considered). Significance threshold was set at $\alpha<0.05$.

Functional analysis of SNPs

Non-synonymous SNPs with a significant effect were analyzed with the Polyphen software (http://genetics.bwh.harvard.edu/pph/) that assesses if the allele variant is probable damaging, possibly damaging or benign on the structure and function of the protein. Polyphen bases the result on analysis of multiple sequence alignments and chemical and structural characteristics of the protein (Ramensky et al. 2002).

Gene expression analysis

mRNA levels of the *ITIH-1*, *-3* and *-4* genes in uterus samples (APICAL and BASAL) from 36 IbxMe F_2 sows classified into two groups according to the NE as follows: LOW (NE<11, n=20), and HIGH (NE>13, n=16) was quantified. Five microliters of a 1/20 dilution of cDNA, synthesized as abovely described, were used as template for real time quantitative PCR (qPCR). The reactions were carried out, in triplicate, following the Power SYBR green PCR Master Mix (Applied Biosystems) manufacturer's instructions. Primer sequences and reaction conditions for qPCR are shown in Supplementary Table 1. All PCR amplifications were performed on an ABI PRISM® 7900 sequence detection system (Applied Biosystems). The porcine *β2-microglobulin* (β2M) gene was used as reference gene (Ferraz et al. 2008) due to its acceptable stability expression evaluated with the Genorm software (M<1.5) and it was used to normalize the cDNA quantity in each real time PCR reaction. PCR efficiencies between target and reference genes were validated for their relative quantification following the $2^{-\Delta\Delta CT}$ method described by Livak and Schmittgen (2001). Resulting qPCR data were $\log_{10}$ transformed and differences between groups were assessed by univariate analysis of variance (ANOVA). For each uterus region, ANOVA with one factor (prolificacy level HIGH or LOW) was used. However, to determine differential expression regarding uterus region, ANOVA with two factors (prolificacy level HIGH or LOW and the uterus region APICAL or BASAL) was performed. ANOVA tests were done with the General Linear Models procedure of the Statistical Analysis System (SAS® Statistics, V 9.1; SAS Institute, Inc., Cary, USA). Correlations were calculated using the *proc corr* procedure of SAS. Significance threshold was set at $\alpha<0.05$.

RESULTS AND DISCUSSION

Polymorphisms

Sequencing of mRNAs from *ITIH-1*, *-3*, and *-4* genes in eleven sows from six different breeds has revealed the existence of 3, 9, and 4 SNPs, respectively, which are segregating in the IbxMe resource population. All three SNPs from *ITIH-1* led to non-synonymous substitutions and none of them was found within the vWA domain. For *ITIH-3*, only four out of nine SNPs from *ITIH-3* determined an amino acid change and six where located in the vWA domain. All four *ITIH-4* SNPs were located in vWA domain but only two of them were non-synonymous. Table 1 shows the position and type of nucleotide and amino acid change for each SNP.

Association studies with prolificacy traits

The resulting linkage map for SSC13 in the IbxMe F_2 population was as follows (distance is indicated in cM): S0076 (0.0) – ITIH1 (18.0) – ITIH3 (18.6) – ITIH4 (19.1) – SWR1008 (27.3) – SW398 (51.6) – SW2440 (74.6) – SW769 (88.0). The position of the *ITIH-1*, *-3* and *-4* genes are in agreement with the pig genome sequence (*Sus scrofa*, Ensembl release 61 - Feb 2011).

Table 2 shows the allelic frequencies of the *ITIH-1*,*-3*, and *-4* SNPs in the IbxMe population as well as the results of the association study of these SNPs with litter size. Significant associations with NBA were found for two SNPs from *ITIH-1* (744 and 2623), four from *ITIH-3* (1332, 1341, 1362 and 1867) and all four SNPs from *ITIH-4* (p<0.05). The favorable alleles from the two *ITIH-1* SNPs and the first three *ITIH-3* SNPs (1332, 1341 and 1362) were present in both parental populations although they were fixed in Meishan sows. These favorable alleles determined an additive effect of almost 0.7. For the remaining SNPs from *ITIH-3* (1867) and the four *ITIH-4* SNPs the favorable alleles were fixed in Meishan sows. In this case, the additive effect was about 0.4. Besides, all *ITIH* SNPs with a significant effect on NBA could explain a proportion of 3% of the phenotypic variance (Table 2). On the other hand, epistatic interactions between *ITIH*

SNPs were tested to assess if interactions between *ITIHs* SNPs have a bigger effect than the individual SNPs. However, no significant epistatic interactions were found between *ITIH* SNPs (data not shown).

Functional analysis of non-synonymous SNPs

The two significant *ITIH-1* SNPs (744 and 2623), one of the four *ITIH-3* significant SNPs (1867) and two out of four significant SNPs from *ITIH-4* (826 and 1027) determined an amino acid substitution. Therefore, we analyzed the translated protein sequences with the Polyphen software (Ramensky et al. 2002), a tool that predicts a possible impact of an amino acid substitution on the structure and function of the encoded protein. All non-synonymous SNPs led to benign amino acid substitutions, thus generating functional proteins.

Marker assisted association test

As shown in Table 3, the haplotypes of *ITIH-1* and *-4* genes remained significant after fitting the QTL as cofactor; whereas, haplotypes of *ITIH-3* gene only reached the suggestive threshold of significance ($p<0.10$). These results indicate that the SSC13 *ITIH* cluster has a slight effect on NBA independently of the QTL located downstream on the chromosome. For *ITIH-1* gene, two haplotypes were found: $ITIH1_A$ (C_{744}-G_{2623}) and $ITIH1_B$ (G_{744}-A_{2623}). Both haplotypes were segregating at equal frequencies in the Iberian boars whereas haplotype $ITIH1_A$ was fixed in Meishan sows. The favorable allele was the one coming from Meishan breed determining an increase of 0.65 piglets per haplotype copy. For *ITIH-3* gene, three haplotypes were found segregating in the IbxMe population: $ITIH3_A$ (T_{1332}-T_{1341}-T_{1362}-A_{1867}), $ITIH3_B$ (C_{1332}-C_{1341}-C_{1362}-G_{1867}), and $ITIH3_C$ (T_{1332}-T_{1341}-T_{1362}-G_{1867}). In this case, haplotype $ITIH3_A$ was fixed in the Meishan population and the other two haplotypes were equally segregating in the Iberian boars. Although no significant differences ($p<0.10$) could be found between these haplotypes, the $ITIH3_A$ and $ITIH3_C$ haplotypes, seemed to determine a higher number of piglets born alive (4.73 and 4.51, respectively) compared to the $ITIH3_B$ haplotype

(4.00). These two haplotypes only differ in the genotype of the last SNP (1867), thereby suggesting that the favorable effect of *ITIH-3* on NBA is mainly determined by the region containing SNPs 1332, 1341, and 1362, which are interestingly located within the vWA domain. Similarly, *ITIH-4* haplotypes ($ITIH4_A$: A_{826}-C_{840}-G_{843}-A_{1027} and $ITIH4_B$: G_{826}-T_{840}-A_{843}-G_{1027}) were alternatively fixed in the Meishan and Iberian populations, respectively, and the favorable effects on NBA came from the haplotype of Meishan origin. As previously mentioned, all four *ITIH-4* SNPs also belong to the vWA domain. The vWA domain is the binding site of such varied molecules as integrins, collagen, proteoglycans and heparin, essential for the synthesis of the extracellular matrix during cumulus expansion (Colombatti and Bonaldo, 1991).

Our results suggest that ten out of sixteen polymorphisms found on the SSC13 *ITIH* cluster have a slight effect on NBA although we cannot discard that these polymorphisms might probably be in linkage disequilibrium with the causal mutation and, therefore, further work should be performed in order to validate this hypothesis. A surprising result, however, was that the effects of the SSC13 *ITIH* cluster was stronger on NBA than on TNB despite the high genetic correlation (r_g~0.9) described between the TNB and NBA traits (Rothschild and Bidanel, 1998). Due to the role that ITIHs play in embryo development and attachment to the uterus and the participation of some of these molecules in inflammatory events we hypothesized that they could have a role in the maintenance of an optimal endometrial environment at the peri-implantation stage. For this reason, we decided to measure the expression levels of the SSC13 *ITIH* cluster in several F_2 IbxMe sows with different reproductive performances.

Expression analyses in high and low prolificacy IbxMe sows

The expression level of the *ITIH-1*, *-3*, and *-4* was determined in APICAL and BASAL uterus of F_2 IbxMe sows classified as of HIGH (n=16) or LOW (n=20) prolificacy based on the number of embryos at 30-32 days of gestation. At 30 days of gestation, the embryo is already attached to the endometrium and between day 30 and 40 of

gestation, the fetal mortality could reach 20% as limited uterine space becomes critical (Ford et al. 2002).

Concerning the level of expression based on prolificacy data by uterus parts, no differences were found between HIGH and LOW prolificacy sows for any of the three *ITIH* genes (Figure 1A and 1B). In BASAL uterus, sows classified as HIGH prolificacy (RQ = 2.29) seemed to have 1.6-fold higher *ITIH-3* expression than LOW prolificacy sows (RQ = 1.45) although no statistical significance could be reached (*P=0.06*, Figure 1B). Unfortunately, the number of F_2 sows classified as of HIGH or LOW prolificacy was limited and could not be increased to have more power in the statistical comparison. For *ITIH-1* and *ITIH-4 mRNAs,* the expression level was similar in HIGH and in LOW prolificacy sows in both uterus parts examined (Figures 1A and 1B). Regarding the comparison between APICAL and BASAL samples, significant differential gene expression was found for *ITIH-3,* whereas no differences were found for *ITIH-1* and *-4. ITIH-3* mRNA was about 1.5 fold higher in APICAL (RQ = 2.62) than in BASAL (RQ = 1.77) samples (*P*=0.01, Figure 1C). Differences in expression profiles between uterus parts is not surprising given that different physiological properties are displayed by different regions of the uterine smooth muscle depending on the phase of the reproductive cycle (Tsirkin, 1986). Interestingly, significant correlations between the expression of *ITIH-1* and *-4* could be determined in both uterus samples whereas *ITIH-3* seemed to follow a different expression pattern (Table 4). These results suggest a different regulation for *ITIH-3* in uterus of pregnant sows. ITIH-3 and -4 proteins have been proposed to be a class of acute phase proteins whose expression is activated as a consequence of an external stimulus such as an infection; whereas the expression of *ITIH-1* gene does not seem to be affected by acute inflammation (Salier et al. 1996; Piñeiro et al. 1999; González-Ramon et al. 2000). Many of the endometrial responses evoked by the developing pig conceptuses during this period of early development and uterine attachment resemble the acute phase response induced during generalized tissue inflammation (Salier et al. 1996; Geisert and Yelich, 1997). Overall, our results suggest that, in pigs, *ITIH-3* might play an important role during early stages of gestation.

It is known that synonymous SNPs do not cause any change in the protein sequence but could affect the gene expression with an effect on the phenotype by influencing splicing accuracy/efficiency or codon-usage bias (Cartegni et al. 2002; Plotkin and Kudla, 2011). To ascertain if the observed polymorphisms could be related with the gene expression level of the *ITIH* SSC13 cluster we have analyzed if the SNPs found in the *ITIH* SSC13 cluster could be related to the gene expression level; however no significant associations could be found.

CONCLUSIONS

Our work has led to the identification of ten SNPs in the *ITIH* SSC13 cluster associated with litter size in IbxMe F_2 population. Furthermore, our results also suggest that higher expression levels of *ITIH-3* are related to prolificacy levels in IbxMe F_2 sows. Although statistical analysis indicate the ten SNPs of the *ITIH* cluster genes as potential markers to be used for selection for litter size in pigs, further analyses in other porcine breeds should be performed to validate their use.

ACKNOWLEDGMENTS

The authors are indebted to L. Varona, M. Arqué, J. Tarrés, M. Fina, and the staff of Nova Genètica, in particular to E. Ramells, F. Márquez, R. Malé, F. Rovira, and I. Riart, for cooperating in the experimental protocol. The authors gratefully acknowledge the contributions of the INRA (France) and the CIA El Dehesón del Encinar (Spain) for providing the purebred Meishan sows and Iberian boars, respectively. This research was funded in part by Project AGL2004-08368-C03-02 and by the Consolider-Ingenio 2010 Program (CSD2007-00036), both from the Spanish Ministry of Science and Innovation. IB is recipient of PIF PhD fellowship from Universitat Autònoma de Barcelona.

REFERENCES

Alves, E., Ovilo, C., Rodríguez, M.C., Silió, L., 2003. Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. Anim. Genet. 34, 319-324.

Baskin, L.C., Pomp, D., Geisert, R.D., 1998. Rapid communication: the porcine Inter-alpha trypsin inhibitor-heavy chain 4 (ITIH4) gene maps to chromosome 13. J. Anim. Sci. 76, 1501-1502.

Bhanumathy, C.D., Tang, Y., Monga, S.P., Katuri, V., Cox, J.A., Mishra, B., Mishra, L., 2002. Itih-4, a serine protease inhibitor regulated in interleukin-6-dependent liver formation: role in liver development and regeneration. Dev. Dyn. 223, 59-69.

Bidanel, J.P., Caritez, J.C., Legault, C., 1989. Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. Genet. Sel. Evol. 21, 507–526.

Bost, F., Diarra-Mehrpour, M., Martin, J.P., 1998. Inter-alpha-trypsin inhibitor proteoglycan family-- a group of proteins binding and stabilizing the extracellular matrix. Eur. J. Biochem. 252, 339-346.

Cartegni, L., Chew, S.L., Krainer, A.R., 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. 3, 285-298.

Chen, L., Wert, S.E., Hendrix, E.M., Russell, P.T., Cannon, M., Larsen, W.J., 1990. Hyaluronic acid synthesis and gap junction endocytosis are necessary for normal expansion of the cumulus mass. Mol. Reprod. Dev. 26, 236-247.

Colombatti, A., Bonaldo, P., 1991. The superfamily of proteins with von Willebrand factor type A-like domains: one theme common to components of extracellular matrix, hemostasis, cellular adhesion, and defense mechanisms. Blood. 77, 2305-2315.

Ferraz, A.L., Ojeda, A., López-Béjar, M., Fernandes, L.T., Castelló, A., Folch, J.M., Pérez-Enciso, M., 2008. Transcriptome architecture across tissues in the pig. BMC Genomics 9, 173.

Ford, S.P., Vonnahme, K.A., Wilson, M.E., 2002. Uterine capacity in the pig reflects a combination of uterine environment and conceptus genotype effects. J. Anim. Sci. 80, E66–E73.

Geisert, R.D., Yelich, J.V., 1997. Regulation of conceptus development and attachment in pigs. J. Reprod. Fertil. Suppl. 52, 133-149.

Geisert, R.D., Ashworth, M.D., Malayer, J.R., 2003. Expression of inter-alpha-trypsin inhibitor heavy chains in endometrium of cyclic and pregnant gilts. Reproduction. 126, 621-627.

González-Ramón, N., Hoebe, K., Alava, M.A., Van Leengoed, L., Piñeiro, M., Carmona, S., Iturralde, M., Lampreave, F., Piñeiro, A., 2000. Pig MAP/ITIH4 and haptoglobin are interleukin-6-dependent acute-phase plasma proteins in porcine primary cultured hepatocytes. Eur. J. Biochem. 267, 1878-1885.

Green, P., Falls, K., Crooks, S., 1990. Documentation for CRI-MAP. Unpublished mimeo (available at: http://www.linkage.Rockefeller.Edu/soft/crimap).

Hettinger, A.M., Allen, M.R., Zhang, B.R., Goad, D.W., Malayer, J.R., Geisert, R.D., 2001. Presence of the acute phase protein, bikunin, in the endometrium of gilts during estrous cycle and early pregnancy. Biol. Reprod. 65, 507-513.

Kadarmideen, H.N., Ali, A.A., Thomson, P.C., Müller, B., Zinsstag, J., 2011. Polymorphisms of the SLC11A1 gene and resistance to bovine tuberculosis in African Zebu cattle. Anim. Genet. (doi:10.1111/j.1365-2052.2011.02203.x).

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods. 25, 402-408.

Noguera, J.L., Rodríguez, C., Varona, L., Tomàs, A., Muñoz, G., Ramírez, O., Barragán, C., Arqué, M., Bidanel, J.P., Amills, M., Ovilo, C., Sánchez, A., 2009. A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. BMC Genomics 10, 636.

Pérez-Enciso, M., Misztal, I., 2004. Qxpak: a versatile mixed model application for genetical genomics and QTL analyses. Bioinformatics 20, 2792-2798.

Piñeiro, M., Alava, M.A., González-Ramón, N., Osada, J., Lasierra, P., Larrad, L., Piñeiro, A., Lampreave, F., 1999. ITIH4 serum concentration increases during acute-phase processes in human patients and is up-regulated by interleukin-6 in hepatocarcinoma HepG2 cells. Biochem. Biophys. Res. Commun. 263, 224-229.

Plotkin, J.B., Kudla, G., 2011. Synonymous but not the same: the causes and consequences of codon bias. Nat. Rev. Genet. 12, 32-42.

Pryce, J.E., Hayes, B.J., Bolormaa, S., Goddard, M.E, 2011. Polymorphic regions affecting human height also control stature in cattle. Genetics. 187, 981-984.

Ramensky, V., Bork, P., Sunyaev, S., 2002. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30, 3894-3900.

Roberts, R.M., Xie, S., Trout, W.E., 1993. Embryo–uterine interactions in pigs during week 2 of pregnancy. J. Reprod. Fertil. Suppl. 48, 171–186.

Ronaghi, M., Pettersson, B., Uhlén, M., Nyrén, P., 1998. PCR-introduced loop structure as primer in DNA sequencing. Biotechniques 25, 876-878, 880-882, 884.

Rothschild, M.F., Bidanel, J.P., 1998. Biology and genetics of reproduction, in: Rothschild, M.F., Ruvinsky, A. (Eds.), The Genetics of the Pig. CAB International, University Press, Cambridge, UK. pp. 313-343.

Salier, J.P., Rouet, P., Raguenez, G., Daveau, M., 1996. The inter-alpha-inhibitor family: from structure to regulation. Biochem. J. 315, 1-9.

Salustri, A., Yanagishita, M., Hascall, V.C., 1989. Synthesis and accumulation of hyaluronic acid and proteoglycans in the mouse cumulus cell-oocyte complex during follicle-stimulating hormone-induced mucification. J. Biol. Chem. 264, 13840-13847.

Scarchilli, L., Camaioni, A., Bottazzi, B., Negri, V., Doni, A., Deban, L., Bastone, A., Salvatori, G., Mantovani, A., Siracusa, G., Salustri, A., 2007. PTX3 interacts with inter-alpha-trypsin inhibitor: implications for hyaluronan organization and cumulus oophorus expansion. J. Biol. Chem. 282, 30161-30170.

Spötter, A., Distl, O., 2006. Genetic approaches to the improvement of fertility traits in the pig. Vet. J. 172, 234-247.

Silió, L., Rodrigáñez, J., Toro, M.A., 2001. La selección de cerdos ibéricos, in: Buxadé, C., Daza, A. (Eds.), Porcino Ibérico: Aspectos Claves. Madrid, pp. 125-149.

Stroband, H.W., Van der Lende, T., 1990. Embryonic and uterine development during early pregnancy in pigs. J. Reprod. Fertil. Suppl. 40, 261-277.

Tsirkin, V.I., 1986. Contractile properties of smooth muscle cells of the horns, cervix and broad ligament of the pig uterus. Fiziol. Zh. SSSR. Im. I. M. Sechenova. 72, 818-829.

Zhao, H., Rothschild, M.F., Fernando, R.L., Dekkers, J.C., 2003. Tests of candidate genes in breed cross populations for QTL mapping in livestock. Mamm. Genome 14, 472-482.

TABLES

Table 1. Polymorphisms of the porcine *ITIH-1*, *-3* and *-4* genes.

Gene	Position	nt change	aa change	Codon	vWA
ITIH-1 ENSSSCT00000012530	744	C/G	Phe/Leu	248	No
	2164	A/G	Arg/Lys	722	No
	2623	G/A	Glu/Lys	875	No
ITIH-3 ENSSSCT00000012532	886	C/T	His/Tyr	296	Yes
	888	C/T	His/Tyr	296	Yes
	909	G/A		303	Yes
	1332	T/C		444	Yes
	1341	T/C		447	Yes
	1362	T/C		454	Yes
	1736	A/C	Lys/Thr	579	No
	1867	A/G	Met/Val	623	No
	1872	T/C		624	No
ITIH-4 ENSSSCT00000012534	826	A/G	Ile/Val	276	Yes
	840	C/T		280	Yes
	843	G/A		281	Yes
	1027	A/G	Lys/Glu	343	Yes

Table 2. Allelic frequencies and results of the association study between *ITIH-1, -3* and *-4* SNPs and prolificacy traits in the IbxMe population.

		Allelic frequencies[2]			TNB (N=858)		NBA (N=858)		
Gene	SNP[1]	Ib (N=3)	Me (N=18)	F_2 (N=218)	a (s.e)[3]	*p*-value[4]	a (s.e)[3]	*p*-value[4]	Effective Size (%)[5]
	744 C>G	0.5	1	0.81	0.57 (0.31)	0.064	0.65 (0.31)	**0.036**	0.33
ITIH1	2164 G>A	1	0.19	0.63	0.15 (0.21)	0.481	0.26 (0.21)	0.220	
	2623 G>A	0.5	1	0.81	0.59 (0.31)	0.054	0.66 (0.31)	**0.030**	0.34
	886 T>C	1	0.31	0.59	0.14 (0.20)	0.476	0.28 (0.20)	0.151	
	888 C>T	1	0.31	0.59	0.14 (0.20)	0.476	0.28 (0.20)	0.151	
	909 A>G	1	0.31	0.59	0.14 (0.20)	0.476	0.28 (0.20)	0.151	
	1332 T>C	0.5	1	0.83	0.60 (0.32)	0.053	0.69 (0.31)	**0.027**	0.37
ITIH3	1341 T>C	0.5	1	0.81	0.60 (0.32)	0.053	0.69 (0.31)	**0.027**	0.37
	1362 T>C	0.5	1	0.81	0.60 (0.32)	0.053	0.69 (0.31)	**0.027**	0.37
	1736 C>A	1	0.28	0.59	0.23 (0.20)	0.553	0.27 (0.20)	0.169	
	1867 A>G	0	1	0.52	0.23 (0.19)	0.219	0.38 (0.19)	**0.044**	0.19
	1872 C>T	0.5	0	0.33	0.01 (0.23)	0.957	0.19 (0.23)	0.418	
	826 A>G	0	1	0.47	0.31 (0.19)	0.106	0.43 (0.19)	**0.022**	0.26
ITIH4	840 C>T	0	1	0.47	0.31 (0.19)	0.106	0.43 (0.19)	**0.022**	0.26
	843 G>A	0	1	0.47	0.31 (0.19)	0.100	0.43 (0.19)	**0.022**	0.26
	1027 A>G	0	1	0.47	0.31 (0.19)	0.106	0.43 (0.19)	**0.022**	0.26

[1] nucleotide position and alleles of the SNP; > indicate the allele that has the major effect.

[2] Allelic frequencies of the favorable allele of each SNP in the F_0 population (Ib = Iberian boars and Me = Meishan sows) and F_2 population.

[3] a = estimation of additive genetic effect; s.e = standard error.

[4] Nominal *p*-value when the effects of *ITIH* polymorphisms are tested.

[5] Effect size $(2pqb^2)/V_A$, where p and q are allele frequencies, b is the additive effect of the SNP solution estimate and V_A is the additive genetic variance.

Table 3. Results of the single association test and the marker assisted association test (MAAT) for ITIH-1, -3 and -4 haplotypes with the number of piglets born alive.

Gene	Haplotype	Haplotype frequencies[1]			Association test (Model 1)		MAAT (Model 2)	
		Ib (N=3)	Me (N=18)	F_2 (N=218)	Haplotype Effect (s.e)[2]	P-value[3]	Haplotype Effect (s.e)[2]	P-value[3]
ITIH-1 n=858	$ITIH1_A$	0.5	1	0.82	4.78 (0.55)	**0.035**	4.78 (0.55)	**0.036**
	$ITIH1_B$	0.5	0	0.18	4.13 (0.58)		4.13 (0.58)	
ITIH-3 n=858	$ITIH3_A$	0	1	0.48	4.73 (0.31)	0.053	4.91 (0.58)	0.063
	$ITIH3_B$	0.5	0	0.17	4.00 (0.40)		4.19 (0.62)	
	$ITIH3_C$	0.5	0	0.34	4.51 (0.33)		4.71 (0.59)	
ITIH-4 n=858	$ITIH4_A$	0	1	0.48	4.81 (0.55)	**0.022**	4.80 (0.55)	**0.028**
	$ITIH4_B$	1	0	0.52	4.37 (0.54)		4.38 (0.54)	

[1]Ib=Iberian, Me=Meishan

[2]s.e= standard error.

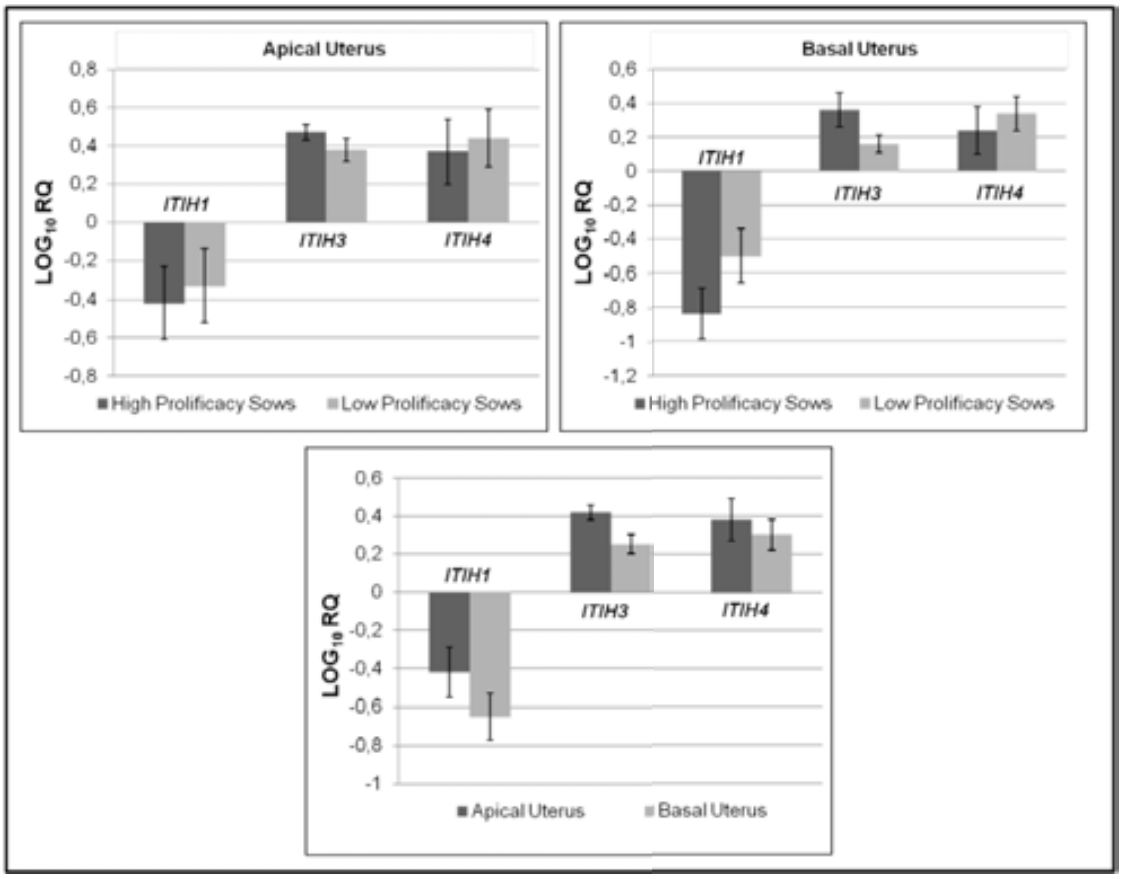
Table 4. Correlation of the expression levels of *ITIH-1*, *-3*, and *-4* genes in apical (above diagonal) and basal (below diagonal) uterus samples.

$Log_{10}RQ$	*ITIH-1*		*ITIH-3*		*ITIH-4*	
	r	*P-Value*	r	*P-Value*	r	*P-Value*
ITIH-1	1		0.34	0.0325	0.84	<0.0001
ITIH-3	0.15	0.3732	1		0.24	0.134
ITIH-4	0.74	<0.0001	0.34	0.0317	1	

r = Pearson Coeficient

FIGURE

Figure 1. Expression of *ITIH-1*, *-3* and *-4* genes between high and low prolificacy sows in (A) apical uterus, (B) basal uterus and (C) between both uterus. Values are represented as log_{10}RQ ± S.E. (Log_{10} relative quantification ± standard error). Relative quantification determines the quantity of a single nucleic acid target sequence within an unknown sample relative to the same sequence within a calibrator sample. Symbol ** significant differences (*P-Value* < 0.05) whereas * suggestive differences (*P-Value* < 0.1) in mRNA expression determined by ANOVA tests.



SUPPLEMENTARY TABLE

Supplementary Table 1. List of primer sequences and reaction conditions for sequencing, typing of polymorphisms and quantitative PCR.

Primer Name[1]	Primer Fw Sequence (5→3')[2]	Primer Rv Sequence (5' → 3')[2]	Conditions	size (bp)	Application
ITIH1-PCR1	gaattctggcctcgtcagg	ccgattgacatcgtaggtca	60ºC, 1,5mM MgCl2, Ecotaq	400	cDNA amplification
ITIH1-PCR2	tcctgctccacatgctctac	aatccacatcatggccga	62ºC, 2,5mM MgCl2, AmpliTaq Gold	595	cDNA amplification
ITIH1-PCR3	ctcaagaacgtccgagatg	ctgatggtcatggaggtcag	60ºC, 2,5mM MgCl2, AmpliTaq Gold	584	cDNA amplification
ITIH1-PCR4	gcagatgtcactggactatcag	atggtcaccaccacctcgt	57ºC, 2,5mM MgCl2, AmpliTaq Gold	555	cDNA amplification
ITIH1-PCR5	accaggcttcactgcg	gaacatagcctctctttaatga	57ºC, 2,5mM MgCl2, AmpliTaq Gold	587	cDNA amplification
ITIH1-PCR6	aacaatcctgctccacatg	aatcacaaaaaccacgttct	54ºC, 1,5mM MgCl2, Ecotaq	257	genomic amplification
ITIH1-PCR6-SNP744[(1)]	gactctgctgaacggggactt		0,5 µM	22	primer extension
ITIH1-PCR7	agtgaatgggcagctcatc	ggtccctccaggagaaca	54ºC, 1,5mM MgCl2, Ecotaq	172	genomic amplification
ITIH1-PCR7-SNP2164[(1)]		gac(tgac)$_3$ttcgcgatccccagccgc	1 µM	35	primer extension
ITIH1-PCR8	tgcaaaaagactacagcaag	tcacgagtgctttgacgta	54ºC, 1,5mM MgCl2, Ecotaq	248	genomic amplification
ITIH1-PCR8-SNP2623[(2)]	gac(tgac)$_5$tatccccggcacg		2 µM	42	primer extension
ITIH1-qPCR	gactccacagaacattacgctg	tcaccaccacctcgtcctg	0,3 µM	95	quantitative PCR
ITIH3-PCR1	atggcatctgctcggt	ccagaggccttgacca	63ºC, 1,5mM MgCl2, Ecotaq	395	cDNA amplification
ITIH3-PCR2	caccaacttcaccttgacca	gtctgtgcaggttgggca	53ºC, 1,5mM MgCl2, Ecotaq	457	cDNA amplification
ITIH3-PCR3	ctcaccaagtccttctca	tacaaggggaacttgcc	55ºC, 2,5mM MgCl2, AmpliTaq Gold	632	cDNA amplification
ITIH3-PCR4	tcccagagaggagcacct	gtttgtccagcagctgc	62ºC, 2,5mM MgCl2, AmpliTaq Gold	562	cDNA amplification
ITIH3-PCR5	acgacctgaccttcaccgag	agcactggcaatgcccagt	62ºC, 2mM MgCl2, Ecotaq	572	cDNA amplification
ITIH3-PCR6	tcacagttaacgggcagatcat	actcagaacaggttggggaca	56ºC, 1,5mM MgCl2, Ecotaq	619	cDNA amplification
ITIH3-PCR7	tagtcaatggctactttgtg	cacgaatattcctcacaa	56ºC, 2,5mM MgCl2, AmpliTaq Gold	503	genomic amplification
ITIH3-PCR7-SNP886 [(3)]	tgacgtcagtggctccatg		1 µM	20	primer extension
ITIH3-PCR7-SNP888 [(3)]		c(tgac)$_2$tacctgctccattttccgacc	1 µM	31	primer extension
ITIH3-PCR7-SNP909 [(3)]		c(tgac)$_4$agaattttgaggagggcatc	1 µM	41	primer extension
ITIH3-PCR8	aatgtgcggaatgcca	aagctgttcatgtcctcatc	62ºC, 1,5mM MgCl2, EcoTaq	1495	genomic amplification
ITIH3-PCR9	aacgcccaaggcaaggagaa	aagccgtggagggggtca	62ºC, 1,5mM MgCl2, EcoTaq	1795	genomic amplification
ITIH3-PCR9-SNP1736	gac(tgac)$_2$caaggcaaggag		3 µM	31	primer extension
ITIH3-PCR9-SNP1867	c(tgac)$_3$tgattttcttcctccct		0,5 µM	41	primer extension
ITIH3-PCR9-SNP1872		gggccaaagctcacctgc	0,5 µM	19	primer extension
ITIH3-PCR10	*ccttgtataacttgggctttgg	caagacaaggcctacctgca	95ºC/60s,58ºC/60s,72ºC/60s , 40cycles	137	pyrosequencing-PCR
ITIH3-PCR10-SNP1872		aactgcaagttggcat			pyrosequencing-probe
ITIH3-qPCR	cacatgggcttctgggacaa	cgatgattcttcaccaccattg	0,3 µM	111	quantitative PCR
ITIH4-PCR1	gccaaaatgaagaccctctc	gtactgctcctgggctgc	59ºC, 1,5mM MgCl2, Ecotaq	321	cDNA amplification
ITIH4-PCR2	cgatggtgtgacctaccca	ctcgtgtcaatgacaaagatca	59ºC, 1,5mM MgCl2, Ecotaq	588	cDNA amplification
ITIH4-PCR3	cggaggtctggtctgcaatac	ccaccaacctcagcagtg	56ºC, 1,5mM MgCl2, Ecotaq	603	cDNA amplification
ITIH4-PCR4	ctgcagctcgaggacttctac	cctgttagtctggatgttctgtc	57ºC, 2,5mM MgCl2, AmpliTaq Gold	575	cDNA amplification
ITIH4-PCR5	atcgagaatggctactttgtgca	cttaattagggcttcccgggt	60ºC, 2,5mM MgCl2, AmpliTaq Gold	343	genomic amplification
ITIH4-PCR6	cctaattaagatcctgggtgacc	cacggccatcagcatcg	60ºC, 2,5mM MgCl2, AmpliTaq Gold	285	genomic amplification
ITIH4-PCR7	cgagaatggctactttgtgca	*tgccagctttcaggaccc	95ºC/30s,60ºC/30s,72ºC/30s, 40cycles	155	pyrosequencing-PCR
ITIH4-PCR7-SNP1027	gcaatacccaagaacg				pyrosequencing-probe
ITIH4-PCR8	*cgtggaggaagccaagagct	gatgggcacatggtctcagg	95ºC/30s,60ºC/30s,72ºC/30s, 40cycles	137	pyrosequencing-PCR
ITIH4-PCR8-SNP1027	ccctgggcatggatt				pyrosequencing-probe
ITIH4-qPCR	cagggggcaactgcacat	aggtatgcccagagtctctccat	0.3 µM	126	quantitative PCR

[1] Numbers between brackets refer to multiplexed reactions.

[2] Biotin-labelled primers are indicated with an asterisk.

3.1.2. Analysis of *MUC4* gene for prolificacy QTL on SSC13

Analysis of porcine *MUC4* gene as a candidate gene for prolificacy QTL on SSC13 in an Iberian x Meishan F_2 population

Balcells, I., Castelló, A., Mercadé, A., Noguera, J.L., Fernández-Rodríguez, A., Sànchez, A. and Tomàs, A.

BMC Genetics (Accepted)

Analysis of porcine *MUC4* gene as a candidate gene for prolificacy QTL on SSC13 in an Iberian x Meishan F_2 population

I. Balcells[1,§], A. Castelló[1], A.Mercadé[2], J.L Noguera[3], A. Fernández-Rodríguez[4], A. Sánchez[1,2] and A. Tomás[5]

[1]Departament de Genètica Animal, Centre de Recerca en Agrigenòmica (CRAG), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[2]Servei Veterinari de Genètica Molecular, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[3]Genètica i Millora Animal, IRTA-Lleida, 25198 Lleida, Spain

[4]Departamento de Mejora Genética Animal, SGIT-INIA, 28040 Madrid, Spain

[5]Program Infection and Immunity, FISIB, 07110 Bunyola, Spain

[§]Correspondence author

ABSTRACT

Background

Reproductive traits, such as prolificacy, are of great interest to the pig industry. Better understanding of their genetic architecture should help to increase the efficiency of pig productivity through the implementation of marker assisted selection (MAS) programmes.

Results

The *Mucin 4* (*MUC4*) gene has been evaluated as a candidate gene for a prolificacy QTL described in an Iberian x Meishan (Ib x Me) F_2 intercross. For association analyses, two previously described SNPs (DQ124298:g.243A>G and DQ124298:g.344A>G) were genotyped in 347 pigs from the Ib x Me population. QTL for the number of piglets born alive (NBA) and for the total number of piglets born (TNB) were confirmed on SSC13 at positions 44 cM and 51 cM, respectively. The *MUC4* gene was successfully located within the confidence intervals of both QTL. Only DQ124298:g.344A>G *MUC4* polymorphism was significantly associated with both NBA and TNB (*P-value* < 0.05) with favourable effects coming from the Meishan origin. *MUC4* expression level was determined in F_2 sows displaying extreme phenotypes for the number of embryos (NE) at 30-32 days of gestation. Differences in the uterine expression of *MUC4* were found between high (NE≥ 13) and low (NE ≤ 11) prolificacy sows. Overall, *MUC4* expression in high prolificacy sows was almost two-fold increased compared with low prolificacy sows.

Conclusions

Our data suggest that *MUC4* could play an important role in the establishment of an optimal uterine environment that would increase embryonic survival during pig gestation.

BACKGROUND

Prolificacy traits have been widely explored during the last decades as a potential tool for increasing efficiency of sow productivity in the pig industry. Genetic improvement programmes have achieved moderate gains in prolificacy related traits owing to their low heritability, late expression in life and sex limitation [1]. Increasing knowledge on the genetic architecture of prolificacy traits would provide new tools to improve the efficiency of genetic selection by implementing marker assisted selection (MAS).

So far a relatively low number of quantitative trait loci (QTL) for prolificacy traits reaching the genome-wide significance level have been identified [2, 3]. The most significant QTL affecting the number of piglets born alive (NBA) and the total number of piglets born (TNB) were described by Noguera *et al.* [3] in the same resource population as used in the present study, an Iberian (Ib) by Meishan (Me) F_2 intercross. A number of epistatic QTL were also detected, thus indicating that the genetic architecture of reproductive traits is built as a complex network of interactions throughout the genome. Some of these epistatic QTL were further confirmed and putative candidate interacting genes were identified [4].

Porcine chromosome 13 (SSC13) harbours the most significant QTL for TNB and NBA [3]. The *Mucin 4* (*MUC4*) gene is located within the confidence interval of prolificacy QTL. Mucins are large heterodimeric glycoproteins commonly located on apical surfaces of many wet-surfaced epithelia that play a key role in the lubrication and protection of the uterine mucosa [5-7]. They have been shown to present anti-adhesive and anti-recognition properties which are necessary to protect the endometrium from the binding and invasion of the trophoectoderm [8, 9]. A role of *MUC4* has been pointed out in rodents and pigs during pregnancy although its expression during the peri-implantational period varies depending on the type of implantation in each species. In mice and rats, which have an invasive type of implantation, *MUC4* expression is downregulated to generate the receptive state for uterine implantation [8-12]. Conversely, in pigs, where a non-invasive epitheliochorial placental attachment takes place, *MUC4* is upregulated in the uterus [13]. A protective role has been suggested for *MUC4* owing to the fact that it is localized on the

endometrium epithelium blocking the access of different substrates to the cell surface [14]. The endometrium is then protected from proteolytic activity of porcine conceptus [13] and from microbial invasion [15] resulting in better uterine conditions for embryo development. In pigs, the disruption of the uterine microenvironment could affect embryo viability which could lead to prenatal mortality rates ranging from 20 to 46 % [16]. The improvement of the uterine microenvironment would increase embryonic survival and, in consequence, the number of piglets born alive.

In humans, polymorphisms in the *MUC4* nucleotide sequence have been significantly associated with the development of endometriosis and endometriosis related infertility [17]. However, no association with implantation failure has been detected [18]. In livestock species, the genetic association of *MUC4* gene variants with reproductive traits has not yet been explored. In pigs, polymorphisms in the *MUC4* gene were shown to be in linkage disequilibrium with susceptibility/resistance to Enterotoxigenic *Escherichia coli* (ETEC) F4ab/ac infection [19].

In the current study, we have examined the porcine *MUC4* gene as a functional and positional candidate gene to explain the prolificacy QTL previously identified on SSC13 in the Ib x Me population [3].

RESULTS

Refinement of SSC13 QTL for NBA and TNB

The resulting linkage map for SSC13 in the Ib x Me population was as follows (distance is indicated in centimorgan (cM)): S0076 (0.0) – ITIH3 (18.7) – SWR1008 (27.5) – MUC4 (39.1) – SW398 (55.8) – SW2440 (78.5) – SW769 (91.5). The position of the *MUC4* gene is in agreement with the pig genome sequence (*Sus scrofa*, Ensembl release 64 - September 2011).

A single QTL scan (model 1) was performed on SSC13 for TNB and NBA (Figure 1). As previously described by Noguera *et al.* [3], two significant QTL for NBA and for TNB were identified at positions 44 cM and 51 cM, respectively (Table 1). Significant additive effects were detected for both QTL which determined an increase of 0.65 ±

0.22 piglets per copy for NBA and 0.51 ± 0.22 for TNB. It is noteworthy that the *MUC4* gene was mapped within the confidence interval (CI) of both QTL.

Candidate gene association analyses

Allele frequencies for DQ124298:g.243A>G and DQ124298:g.344A>G SNPs in the Ib x Me population are shown in Table 2. Associations between *MUC4* polymorphisms and reproductive traits were tested with a standard animal model (model 2). No significant associations were found for the DQ124298:g.243A>G SNP and the reproductive traits recorded (data not shown). For the DQ124298:g.344A>G SNP, the results supported significant additive effects between this SNP and NBA and TNB traits (*P-value* < 0.05, Table 1). The additive substitution effect for DQ124298:g.344G SNP was estimated to be 0.74 ± 0.27 for NBA (*P-value* = 0.006) and 0.57 ± 0.27 for TNB (*P-value* = 0.037). In both cases, the G allele coming from the Me breed had a favourable effect. Note that DQ124298:g.344A>G SNP had a larger effect on NBA than on TNB (*P-value* = 0.006 for NBA and *P-value* = 0.037 for TNB).

The results obtained with association studies must be interpreted with caution due to the fact that the extensive linkage disequilibrium described in F_2 crosses can bias the estimations. For this reason, in order to detect possible false positives obtained in association studies with DQ124298:g.344A>G SNP, a marker assisted association test (MAAT) including the QTL effect in the association test [20] was performed with model 3. Results from the MAAT are summarized in Table 1. The DQ124298:g.344A>G SNP effect on NBA and TNB remained significant when the QTL effect was considered which is in agreement with the association studies analyses. Nevertheless, the significance of the QTL disappeared when the DQ124298:g.344A>G SNP was included in the model. MAAT shows that the DQ124298:g.344A>G SNP genotype explains better the effects on NBA and TNB and confirms the association detected with model 2.

Expression analysis of the porcine *MUC4* gene

In order to determine whether *MUC4* expression could affect prolificacy related traits, we analysed the uterine expression profile of porcine *MUC4* in sows that differed in the number of embryos (NE) at 30-32 days of gestation using qPCR. At this time of

gestation, the embryo is already attached to the endometrium and the foetal survival rate will be an indication of the final litter size [1]. For this reason, NE was measured as an estimation of prolificacy. Results showed that mRNA expression levels of *MUC4* gene were suggestively greater in high (NE ≥ 13, n = 16, mean relative expression = 7.22) than in low (NE ≤ 11, n = 20, mean relative expression = 3.63) prolificacy sows (*P-value* = 0.07, Figure 2), reaching almost a two-fold increase in the high prolificacy group.

DISCUSSION

Statistical association between the DQ124298:g.344A>G mutation in the porcine *MUC4* gene and prolificacy related traits has been reported. No association was found, however, between the DQ124298:g.243A>G SNP and the prolificacy related traits although this SNP is located 100bp upstream of the DQ124298:g.344A>G mutation. The difference in association analysis detected between both *MUC4* SNPs indicates that they are not in linkage disequilibrium. The *MUC4* gene successfully mapped within the confidence interval of the SSC13 QTL affecting NBA and TNB traits previously described by Noguera *et al.* [3] using the same resource population. The favourable effect of DQ124298:g.344A>G SNP was inherited from the Me maternal breed, as expected, since Me is more prolific than Ib. The effect of the DQ124298:g.344A>G *MUC4* SNP was stronger for NBA than for TNB suggesting that the *MUC4* gene could be related to the embryonic survival in the uterus. This hypothesis is supported by the differential expression of the *MUC4* gene found in the uterus at 30-32 days of gestation, where the expression was two-fold higher in high than in low prolificacy sows. Nevertheless, further studies must be performed in other unrelated porcine populations to validate the association which was obtained in this study.

MUC4 has been identified as a potential regulator of placentation in pigs [21]. In porcine uterine surface epithelium, *MUC4* gene expression increases during the first 18 days of gestation [13] whereas in rats a reduction in *MUC4* gene expression in the uterus is associated with the period of implantation [12]. Differences in *MUC4* expression profiles could be explained by the different placentation types of rodents

and pigs. Pigs have a non-invasive type of placentation and MUC4 is thought to play a role in protecting the uterus from erosion by the embryo [13, 15]. MUC4 and MUC1 are the major mucin proteins expressed in the endometrial epithelium [22, 23]. Human *MUC1* and *MUC4* present highly polymorphic sites with a variable number of tandem repeats (VNTRs). *MUC1* VNTR variants have been related to alterations in both T-antigen presentation and in the local immune response in cancer [17, 24, 25]. These results suggest that mucins may play a role in the immunological processes that take place during the implantation period essential to ensure the correct establishment of maternal-foetal tolerance [26].

CONCLUSIONS

MUC4 polymorphism DQ124298:g.344A>G is associated with litter size in the Ib x Me population and, moreover, *MUC4* is differentially expressed regarding the number of embryos in uterus at 30 days of gestation. These results suggest that *MUC4* may participate in the establishment of an optimal uterine environment essential for adequate embryo development during the early stages of gestation and increase litter size in pigs.

METHODS

Animal material and phenotypic measurements

An F_2 population was generated by crossing 3 Ib males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) with 18 Me females (Domaine du Magneraud, INRA, France). A total of 8 boars and 97 sows from the F_1 generation were mated to obtain the F_2 progeny in the Nova Genètica S.A experimental farm (Lleida, Spain). All animals were obtained according to the European animal experimentation ethics law approved by the Ethical and Care Committee at IRTA.

Measurements of sow reproduction traits including TNB and NBA were recorded in 255 F_2 sows during 4 successive parities. In the fifth parity, sows were slaughtered at

30-32 days of gestation, when the uterus samples were recollected and the number of embryos (NE) was recorded.

Genotyping of *MUC4* polymorphisms

Two *MUC4* polymorphisms described by [19], two A-to-G substitution polymorphisms at positions 243 and 344 of intron 17 (GenBank accesion number DQ124298), were genotyped in our Ib x Me population by pyrosequencing [27] in a PSQ HS 96 system (Pyrosequencing AB, Uppsala, Sweden). Two PCR amplifications, one for each SNP, were carried out in a 25 µl total volume that included 1.5 mM of $MgCl_2$, 200 µM of each dNTP, 0.75 U of TaqGold DNA polymerase (Applied Biosystems), 320 nM of each primer (Table 3) and 40 ng of DNA. The thermal profile was 95 ºC for 10 min, 40 cycles at 94 ºC for 30 s, 58 ºC for 45 s and 72 ºC for 1 min and a final extension step of 15 min at 72 ºC. A multiplex pyrosequencing reaction was performed with 5 µl of each PCR with the primers described in Table 3.

QTL and association analyses

The linkage map of porcine chromosome 13 (SSC13) was constructed with the *Build* option of CRIMAP 2.4 software [28]. Overall, seven markers were used: five microsatellites that had been previously described by [3], one SNP at the *ITIH3* gene (dbSNP accession number ss315834911) and the DQ124298:g.344A>G *MUC4* SNP described by [19].

Three models were used to analyse the Ib x Me F_2 population data: (1) a QTL model, (2) an association model and (3) a QTL + association model to perform the marker assisted association test (MAAT) proposed by [20].

First, one dimensional QTL mapping was performed with model 1.

$$y_{ijk} = H_i + O_j + u_k + p_k + C_a a + e_{ijkl} \quad \text{(model 1)}$$

where y_{ijk} was the phenotypic data vector for NBA or TNB; H_i and O_j were the fixed effects for year-season and parturition order, respectively; u_k was the random polygenic effect of each individual; p_k was the environmental permanent effect of the sow; a was the QTL additive effect; C_a was the probability of the individual being

homozygous for Ib alleles minus the probability of being homozygous for Me alleles at the QTL position of interest; and e_{ijkl} was the random residual term. The dominance effect was not included in the model because the likelihood ratio test performed indicated that a model with only additive QTL effect fitted better.

Second, association analyses were performed with the *MUC4* SNPs (DQ124298:g.243A>G and DQ124298:g.344A>G) with a standard animal model (model 2).

$$y_{ijk} = H_i + O_j + u_k + p_k + \sum_k \lambda_{ik}\, a_k + e_{ijkl} \qquad \text{(model 2)}$$

where y_{ijk} was the vector containing the phenotypic data for NBA and TNB and λ_{ik} was a variable that indicated the number of copies (0, 1, or 2) of the k^{th} allele presented by each individual.

Finally, a combined QTL + association model (model 3) was used to consider the extensive linkage disequilibrium present in F_2 population. Only the DQ124298:g.344A>G SNP was considered because significant results were obtained in model 2.

$$y_{ijk} = H_i + O_j + u_k + p_k + C_a a + \sum_k \lambda_{ik}\, a_k + e_{ijkl} \qquad \text{(model 3)}$$

All analyses were performed with Qxpack software [29]. QTL scans were analysed every cM and nominal *P-value*s were calculated with the maximum likelihood ratio test, assuming a χ^2 distribution of the likelihood ratio test with degrees of freedom calculated as the difference between the number of parameters in the alternative and in the null models.

Expression analysis of the *MUC4* gene

Uterine expression of the *MUC4* gene was measured by reverse transcription quantitative real time PCR (RT-qPCR). Total RNA from uterus samples was extracted by means of the RiboPure™ kit (Ambion, Applied Biosystems). One microgram of total RNA in 40 µl reaction was reverse transcribed with the High Capacity cDNA Transcription Kit (Applied Biosystems). Primers for *MUC4* (Table 3) were designed with Primer Express® 2.0 software (Applied Biosystems, Warrington, UK). *Sus scrofa*

hypoxanthine phosphoribosyltransferase 1 (*HPRT1*) was used as a reference gene for normalization [for primer sequences see 30]. qPCR reactions were performed in triplicate in a 20 µl final volume including 2X FastStart SYBR Green Master (Roche), 0.3 µM of each primer and 5 µl of the cDNA diluted twenty times on an ABI PRISM® 7900HT sequence detection system (Applied Biosystems, Warrington, UK). Thermal conditions were 95 ºC for 10 min followed by 40 cycles of 95 ºC for 15 s and 60 ºC for 1 min. Dissociation curve analyses were performed in order to detect unspecific amplifications. PCR efficiencies of *MUC4* and *HPRT1* genes were calculated to validate the use of the $2^{-\Delta\Delta Ct}$ method [31]. The *MUC4* gene expression level was measured individually in 36 Ib x Me F_2 sows from 32 different litters that were classified into two groups according to the NE at slaughter (30-32 days of pregnancy): low (NE≤ 11, n = 20) and high (NE≥ 13, n = 16). Gene expression data were $\log_{10}$ transformed and analysed by a t-test using the Statistical Analysis System (Statistics, V 9.1.3; SAS Institute, Inc., Cary, NC). The significance threshold was set at $\alpha < 0.05$.

AUTHORS' CONTRIBUTIONS

IB carried out qPCR experiments, analysed and interpreted the data and prepared the manuscript. AC performed the genotyping task and revised the manuscript. AM designed the genotyping protocol. JLN participated in the design of the study and coordinated it. AF participated in sample collection and helped with the critical revision of the manuscript. AS conceived the study, participated in the design and the supervision of the study and revised the draft. AT supervised the study, helped to draft the manuscript and revised it. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors are indebted to L. Varona, M. Arqué, J. Tarrés, M. Fina, and the staff of Nova Genètica, particularly E. Ramells, F. Márquez, R. Malé, F. Rovira, and I. Riart, for cooperating in the experimental protocol. The authors gratefully acknowledge the contributions of the INRA (France) and the CIA El Dehesón del Encinar (Spain) in

providing the purebred Meishan sows and Iberian boars, respectively. The authors also thank J. Casellas, V. Martínez and S. Cirera for critical comments on the manuscript. This research was funded in part by Project AGL2004-08368-C03 and by the Consolider-Ingenio 2010 Programme (CSD2007-00036), both from the Spanish Ministry of Science and Innovation. IB is the recipient of a PIF PhD fellowship from Universitat Autònoma de Barcelona.

REFERENCES

1. Spotter A, Distl O: Genetic approaches to the improvement of fertility traits in the pig. Vet J 2006, 172:234-247

2. Tribout T, Iannuccelli N, Druet T, Gilbert H, Riquet J, Gueblez R, Mercat MJ, Bidanel JP, Milan D, Le Roy P: Detection of quantitative trait loci for reproduction and production traits in Large White and French Landrace pig populations. Genet Sel Evol 2008, 40:61-78

3. Noguera JL, Rodriguez C, Varona L, Tomas A, Munoz G, Ramirez O, Barragan C, Arque M, Bidanel JP, Amills M, Ovilo C, Sanchez A: A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. BMC Genomics 2009, 10:636

4. Fernandez-Rodriguez A, Rodriguez C, Varona L, Balcells I, Noguera JL, Ovilo C, Fernandez AI: Analysis of candidate genes underlying two epistatic quantitative trait loci on SSC12 affecting litter size in pig. Anim Genet 2010, 41:73-80

5. Moniaux N, Nollet S, Porchet N, Degand P, Laine A, Aubert JP: Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. Biochem J 1999, 338:325-333

6. Lagow E, DeSouza MM, Carson DD: Mammalian reproductive tract mucins. Hum Reprod Update 1999, 5:280-292

7. Govindarajan B, Gipson IK: Membrane-tethered mucins have multiple functions on the ocular surface. Exp Eye Res 2010, 90:655-663

8. McNeer RR, Carraway CA, Fregien NL, Carraway KL: Characterization of the expression and steroid hormone control of sialomucin complex in the rat uterus: implications for uterine receptivity. J Cell Physiol 1998, 176:110-119

9. McNeer RR, Huang D, Fregien NL, Carraway KL: Sialomucin complex in the rat respiratory tract: a model for its role in epithelial protection. Biochem J 1998, 330:737-744

10. Idris N, Carraway KL: Sialomucin complex (Muc4) expression in the rat female reproductive tract. Biol Reprod 1999, 61:1431-1438

11. Idris N, Carraway KL: Regulation of sialomucin complex/Muc4 expression in rat uterine luminal epithelial cells by transforming growth factor-beta: implications for blastocyst implantation. J Cell Physiol 2000, 185:310-316

12. Carraway KL, Idris N: Regulation of sialomucin complex/Muc4 in the female rat reproductive tract. Biochem Soc Trans 2001, 29:162-166

13. Ferrell AD, Malayer JR, Carraway KL, Geisert RD: Sialomucin complex (Muc4) expression in porcine endometrium during the oestrous cycle and early pregnancy. Reprod Domest Anim 2003, 38:63-65

14. Komatsu M, Carraway CA, Fregien NL, Carraway KL: Reversible disruption of cell-matrix and cell-cell interactions by overexpression of sialomucin complex. J Biol Chem 1997, 272:33245-33254

15. Carraway KL, Perez A, Idris N, Jepson S, Arango M, Komatsu M, Haq B, Price-Schiavi SA, Zhang J, Carraway CA: Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. Prog Nucleic Acid Res Mol Biol 2002, 71:149-185

16. Ross JW, Ashworth MD, Stein DR, Couture OP, Tuggle CK, Geisert RD: Identification of differential gene expression during porcine conceptus rapid trophoblastic elongation and attachment to uterine luminal epithelium. Physiol Genomics 2009, 36:140-148

17. Chang CY, Chang HW, Chen CM, Lin CY, Chen CP, Lai CH, Lin WY, Liu HP, Sheu JJ, Tsai FJ: MUC4 gene polymorphisms associate with endometriosis development and endometriosis-related infertility. BMC Med 2011, 9:19

18. Koscinski I, Viville S, Porchet N, Bernigaud A, Escande F, Defossez A, Buisine MP: MUC4 gene polymorphism and expression in women with implantation failure. Hum Reprod 2006, 21:2238-2245

19. Peng QL, Ren J, Yan XM, Huang X, Tang H, Wang YZ, Zhang B, Huang LS: The g.243A>G mutation in intron 17 of MUC4 is significantly associated with susceptibility/resistance to ETEC F4ab/ac infection in pigs. Anim Genet 2007, 38:397-400

20. Zhao H, Rothschild MF, Fernando RL, Dekkers JC: Tests of candidate genes in breed cross populations for QTL mapping in livestock. Mamm Genome 2003, 14:472-482

21. Ostrup E, Bauersachs S, Blum H, Wolf E, Hyttel P: Differential endometrial gene expression in pregnant and nonpregnant sows. Biol Reprod 2010, 83:277-285

22. Gollub EG, Goswami S, Kouba D, Marom Z: Regulation of mucin gene expression in secretory epithelial cells. Biochem Biophys Res Commun 1993, 197:667-673

23. Audie JP, Tetaert D, Pigny P, Buisine MP, Janin A, Aubert JP, Porchet N, Boersma A: Mucin gene expression in the human endocervix. Hum Reprod 1995, 10:98-102

24. Fowler JC, Teixeira AS, Vinall LE, Swallow DM: Hypervariability of the membrane-associated mucin and cancer marker MUC1. Hum Genet 2003, 113:473-479

25. Santos-Silva F, Fonseca A, Caffrey T, Carvalho F, Mesquita P, Reis C, Almeida R, David L, Hollingsworth MA: Thomsen-Friedenreich antigen expression in gastric carcinomas is associated with MUC1 mucin VNTR polymorphism. Glycobiology 2005, 15:511-517

26. Guerin LR, Prins JR, Robertson SA: Regulatory T-cells and immune tolerance in pregnancy: a new target for infertility treatment? Hum Reprod Update 2009, 15:517-535

27. Ronaghi M, Pettersson B, Uhlen M, Nyren P: PCR-introduced loop structure as primer in DNA sequencing. BioTechniques 1998, 25:876-8, 880-2, 884

28. Green P, Falls K, Crooks S: Documentation of CRI-MAP [http://compgen.rutgers.edu/old/multimap/crimap/index.html]

29. Perez-Enciso M, Misztal I: Qxpak: a versatile mixed model application for genetical genomics and QTL analyses. Bioinformatics 2004, 20:2792-2798

30. Fernandes LT, Tomas A, Bensaid A, Perez-Enciso M, Sibila M, Sanchez A, Segales J: Exploratory study on the transcriptional profile of pigs subclinically infected with porcine circovirus type 2. Anim Biotechnol 2009, 20:96-109

31. Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001, 25:402-408

TABLES

Table 1. Results of QTL analyses, association tests and marker assisted association tests for prolificacy traits.

	QTL model (model 1)			*MUC4* association (model 2)		QTL + *MUC4* association model (model 3)[1]			
Trait	Pos. (c.i)[2]	*P*-QTL	a_{QTL} (SE)[3]	*P*-MUC4	a_{MUC4} (SE)[4]	Pos.[5]	*P*-QTL + MUC4 (DQ124298:g.344A>G)	*P*-QTL	*P*-MUC4 (DQ124298:g.344A>G)
NBA	44 (22-77)	0.003	0.65 (0.22)	0.006	0.74 (0.27)	63	0.003	0.062	0.007
TNB	51 (33-74)	0.021	0.51 (0.22)	0.037	0.57 (0.27)	64	0.019	0.061	0.039

[1]*P-values* obtained with the model that includes the QTL and the DQ124298:g.344A>G *MUC4* SNP effects; each effect(s) was tested by removing it from the null model; [2]QTL position (in cM); c.i = confidence interval; [3]QTL additive effect; SE = standard error; [4]DQ124298:g.344A>G *MUC4* SNP additive effect; [5]QTL position in cM when corrected by the DQ124298:g.344A>G *MUC4* SNP effect.

Table 2. Allelic frequencies of the DQ124298:g.243A>G and DQ124298:g.344A>G *MUC4* polymorphisms in the Me x Ib F_2 population.

		DQ124298:g.243			DQ124298:g.344		
		n	A	G	n	A	G
F_0	♀ Meishan	18	0	1	18	0	1
	♂ Iberian	3	0.50	0.50	3	0.50	0.50
	F_1	120	0.32	0.68	123	0.18	0.82
	F_2	206	0.20	0.80	202	0.29	0.71

Table 3. List of primer sequences used for typing *MUC4* polymorphisms and quantitative PCR.

Application	SNP/Gene	Primer Sequence (5' →3')	Length (bp)
Pyrosequencing PCR	DQ124298:g.243A>G	F: *tggtgctacccccagatttg	195
		R: gttgtgtccacccccttacccttat	
		P: gtcccctctcccaggta	
	DQ124298:g.344A>G	F: *gtggccctcagtcactagagt	258
		R: cgaagttgtgaaaggaagacag	
		P: ttggggttggggcag	
qPCR	*MUC4*	F: atgggcttctccagtggagat	66
		R: tctcccacactggctgcaa	

* 5´ Biotin labelled, F forward primer, R reverse primer, P pyrosequencing primer

FIGURES

Figure 1. QTL profiles on SSC13 for NBA and TNB showing the likelihood ratio test statistic. The horizontal lines set at 3.84 and 6.63 show the 0.05 and the 0.01 significance levels respectively.

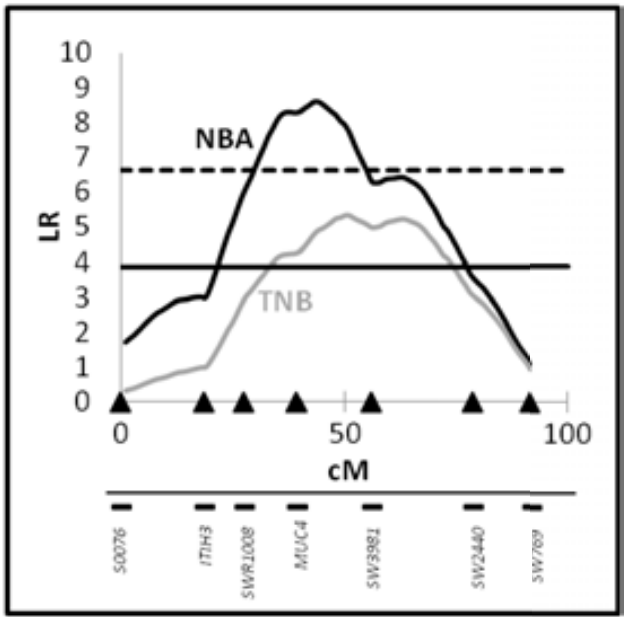
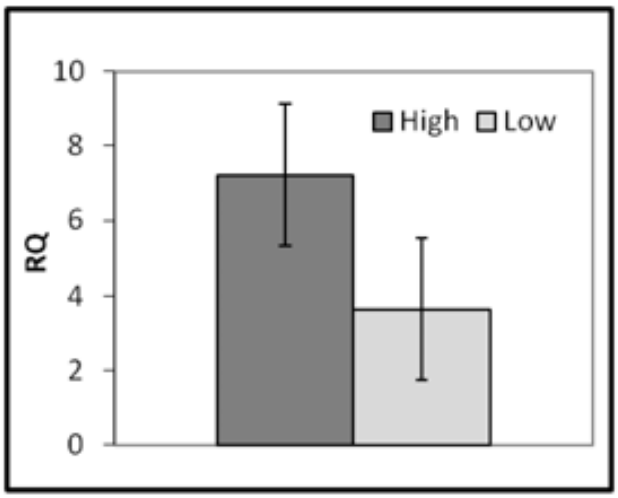


Figure 2. Relative quantification (RQ) of the porcine *MUC4* gene in the uterus QTL profiles on SSC13 for NBA and TNB showing the likelihood ratio test statistic. *MUC4* expression in the uterus was measured in 36 F_2 sows that were classified into two groups according to the number of embryos (NE) at the sacrifice day (30-32 days of gestation): low (NE ≤ 11) and high (NE ≥ 13).



3.2. TRANSCRIPTOME ANALYSES

3.2.1. Analysis of mRNA and miRNA in uterus of pregnant IbxMe F_2 sows regarding prolificacy levels.

Deciphering the uterine transcriptome and microRNAome of high and low prolificacy IbxMe F_2 sows

Balcells, I., et al.

Manuscript in preparation

Deciphering the uterine transcriptome and microRNAome of high and low prolificacy IbxMe F_2 sows

Balcells I[1,*] et al. [2,3,4,5,6,7] (manuscript in preparation)

[1]Departament de Genètica Animal, Centre de Recerca en Agrigenòmica (CRAG), Universitat Autònoma de Barcelona, Bellaterra, Spain

[2]Department of Basic Animal and Veterinary Sciences, Division of Genetics and Bioinformatics, University of Copenhagen, Copenhagen, Denmark

[4]Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona, Bellaterra, Spain

[5]Departamento de Mejora Genética Animal, SGIT-INIA, Madrid, Spain

[6]Departament de Genètica i Millora Animal, IRTA, Lleida, Spain

[7]Program Infection and Immunity, FISIB, Bunyola, Spain

[*]Correspondence author

INTRODUCTION

Understanding the biological basis of reproductive function is essential for the improvement of reproductive efficiency in pigs. Selection programs towards the improvement of prolificacy traits, such as litter size, ovulation rate and embryo survival, have typically yielded low-to-moderate increases due to the low heritability, late expression in life and sex-limited expression displayed by these traits. The discovery of genes and/or genomic regions associated with the phenotypic expression of such traits has been a major goal for the scientific community during the last decades. A number of studies aimed at identifying quantitative trait loci (QTL) segregating in experimental and commercial pig populations has led to the discovery of several chromosomal regions with effects on the expression of reproductive traits (Rothschild *et al.*, 2007). However, the lack of identification of the true causal gene/mutation due to the large disequilibrium linkage blocks present in the genome of livestock species seems to be one of the main limitations of this type of approaches. In particular, linkage analyses for litter size have been scarce and scarcely significant (Onteru *et al.*, 2009,Caetano *et al.*, 2004). Only one study performed by our group could detect significant QTL for litter size traits on porcine chromosomes 13 (SSC13) and SSC17 (Noguera *et al.*, 2009). In addition, a number of epistatic QTL were also detected showing that the genetic architecture of prolificacy in pigs is a complex system of interactions throughout the genome.

In the recent years, the implementation of large-scale gene expression analysis has raised as a useful complementary approach to identify those genes with a function in the biology of reproduction, although the number of transcriptome analyses in porcine species is still low. There are two main strategies for the characterisation of the gene expression on a genome-wide scale: the microarray technology (Schena *et al.*, 1995) and the high-throughput sequencing (HTS) of transcripts (Mardis, 2008, Velculescu *et al.*, 1995). The first genome-wide porcine expression studies were conducted using human microarrays (Medhora *et al.*, 2002, Moody *et al.*, 2002). Since then, the number of microarrays targeting porcine-specific probes has increased markedly (Tuggle *et al.*, 2007). So far, several studies of gene expression analyses in porcine reproductive tissues have been conducted (Onteru *et al.*,

2009,Caetano *et al.*, 2004,Gladney *et al.*, 2004). Most of them have been mainly focused on the characterization of the normal physiological processes in porcine reproductive tissues such as folliculogenesis, embryo development, pregnancy recognition and maintenance, and embryo implantation (Caetano *et al.*, 2004,Gladney *et al.*, 2004,Alminana *et al.*, 2012,Blomberg le *et al.*, 2010,Ostrup *et al.*, 2010,Ross *et al.*, 2007).

However, the characterization of differential gene expression due to genetic variation has been scarce, with only two studies tackling the differential gene expression displayed by European and Chinese porcine breeds in the ovary (Sun *et al.*, 2011,Fernandez-Rodriguez *et al.*, 2011). The main limitations of microarrays studies in pigs have laid on the incomplete knowledge of the porcine genome at the moment of probe design. These limitations were circumvented by the development of HTS technologies, which have allowed the sequencing of the whole porcine transcriptome. At the moment, however, only few RNAseq studies have been conducted in porcine reproductive tissues (Esteve-Codina *et al.*, 2011,Uenishi *et al.*, 2007).

Another layer of complexity in the molecular characterisation of the biology of reproduction has been added by the discovery of the small non-coding RNAs (ncRNAs). The most widely studied class of small ncRNAs are the microRNAs (miRNAs) (21-23 nucleotides) which have been shown to be involved in the regulation of gene expression in many biological processes including reproduction (Nothnick, 2012,Hawkins *et al.*, 2011). Up to date the number of characterised porcine miRNAs included in the miRBase database v18 (November 2011) is of about 257 mature miRNAs, but it has been predicted by *in silico* analysis of the porcine genome sequence that this number could raise up to 1000 miRNAs (Seemann *et al.*, 2007). Several miRNAs have been identified to play a key role in the regulation of steroid hormones essential for reproduction (Ishibashi *et al.*, 2012,Xu *et al.*, 2011). In pigs, miR-27a is differentially expressed in ovaries of Chinese Erhualian and Large White breeds and one SNP has been associated with litter size (Lei *et al.*, 2011).

In the present work, we have characterised the uterine transcriptome and microRNAome of Iberian (Ib) x Meishan (Me) F_2 sows at 30-32 days of pregnancy displaying extreme phenotypes regarding the number of embryos at 30-32 days of gestation.

MATERIAL AND METHODS

Animal Material

Animals used for this study come from an F_2 experimental cross between Ib and Me breeds, which was thoroughly described in (Noguera *et al.*, 2009). Briefly, 18 Me sows were mated to 3 Ib boars (Guadyerbas line) to obtain 255 F_2 sows. Reproductive performance of four consecutive parities was recorded for each sow. In the fifth gestation, sows were slaughtered at day 30-32 of pregnancy and the number of embryos (NE) attached to the uterus were annotated. Sows were classified as high or low prolificacy based on the NE. Uterus sample from the apical part were collected at slaughter, snap frozen in liquid nitrogen and stored at -80ºC until usage. All animal procedures were carried out according to Spanish and European animal experimentation ethics law and approved by the institutional animal ethics committee of IRTA.

Microarray Analysis

Total RNA was extracted from uterus of 14 IbxMe F_2 sows divided into two prolificacy groups according the NE: high (n = 5; NE $\geq$ 14) and low (n = 9; NE $\leq$ 11) prolificacy sows. The total RNA extraction was performed with the RiboPure™ kit (Ambion, Austin, USA), following the manufacturer's instructions. Total RNA quality was assessed by capillary electrophoresis on an Agilent 2100 Bioanalyzer using the RNA Nano 6000 Labchip kit (Agilent Technologies, Palo Alto, USA). NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA) was used in order to quantify the total RNA. Only total RNA with RNA integrity number (RIN) $\geq$ 7

was used. Samples were hybridized to the Affymetrix 24K Genechip® Porcine Genome Array (Affymetrix, Santa Clara, CA, USA) following the standard Affymetrix one-cycle protocol. Reverse transcription, RNA labelling, cRNA amplification, hybridization and scanning procedures were conducted at the Affymetrix facilities available at the Institut de Recerca Hospital Universitari Vall d'Hebron, Barcelona, Spain. Briefly, biotin-labelled cRNA was produced from 5 µg of total RNA using an Affymetrix "one-way" labelling kit (Affymetrix, Expression Analysis Technical Manual). Labelled cRNA was fractioned and hybridised with the GeneChip® Porcine Genome Array following the manufacturer's instructions. Subsequently, the arrays were stained with a GeneChip Fluidics Station 450 using the standard fluidics protocol and were scanned with an Affymetrix GeneChip Scanner 3000 (Affymetrix) by using the GeneChip Operating Software (GCOS) to produce a *.CEL file for further data processing.

Microarray data analysis

Raw data and statistical analyses were performed with Bioconductor (Gentleman *et al.*, 2004) implemented in R 2.6.0 (http://cran.r-project.org/). Data quality was assessed by the QC function implemented in the *simpleaffy* package (Wilson and Miller, 2005). The Robust Multichip Average (RMA, (Irizarry *et al.*, 2003) methodology was used for array normalization. The Empirical Bayes t-test statistic implemented in the *limma* package (Smyth, 2004) was used to determine differential gene expression between high and low prolificacy sows. The threshold of significance was set to a false discovery rate (FDR, (BENJAMINI and HOCHBERG, 1995)) of 0.05 and a minimum fold change (FC) of 2. Hierarchical clustering was performed with Cluster 3.0 and Java TreeView 1.1 software, using the uncentered correlation coefficient and the average linkage method. Probes were annotated based on the chip annotation provided by Affymetrix (NetAffx), (Tsai *et al.*, 2006) and the annotation of Iowa State University (Chris Tuggle, personal communication). However, some of the probes were not coincident between different sources and were, therefore, validated by screening the probe nucleotide sequence available at

NetAffx with the nr and EST databases available at NCBI using the Basic Local Alignment Search Tool (BLAST, http://blast.ncbi.nlm.noh.gov/Blast.cgi). Those probes that could not be assigned to a known gene were not used for functional analyses. The Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/home.jsp) was used for assessing functional profiles of genes based on the Biological Processes (BP) category of Gene Ontology (GO).

Quantitative real-time PCR (RT-qPCR)

To confirm microarray data, quantitative real-time PCR (RT-qPCR) was done for 4 genes (*BMP2*, *CALB3*, *PKP2* and *PTGIS*) for uterus samples. Firstly, the expression levels of these genes were measured by using the 14 IbxMe F_2 sows used in microarray study. Afterward, the sample size was expanded to n = 36; 20 from low prolificacy (NE ≤ 11) and 16 from high prolificacy (NE ≥ 13)). Only total RNA with RIN ≥ 7 was used. The *hypoxantine phosphoribosyltransferase* (*HPRT1*) gene was used as a reference housekeeping gene (Fernandes *et al.*, 2009). Porcine specific primers were designed with the Primer Express 2.0 software (Applied Biosystems, Warrington, UK). Primer sequences are shown in Table 1. cDNA synthesis was performed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster city, USA) using one microgram of total RNA. RT-qPCR was performed in triplicate in a 20 µL final volume reaction containing 5 µL of a 1:20 dilution of the cDNA, 300 nM of each primer (Table 1) and 10 µL of 1X Power SYBR green PCR Master Mix (Roche, Indianapolis, USA) on an 7900HT Sequence Detection System (Applied Biosystems). Thermal profile consisted of a denaturalization step at 95 ºC for 10 min followed by 40 cycles at 95 ºC/15 s and 60 ºC/1 min. PCR efficiencies between target and reference genes were validated for their relative quantification following the $2^{-\Delta\Delta Ct}$ method described by (Livak and Schmittgen, 2001). Resulting RT-qPCR data were Log_{10} transformed and analyzed, on a gene-by-gene basis, with the proc GLM method of SAS software (Statistics, V 9.1; SAS Institute, Inc.,

Cary, NC, USA) following the models used for microarray data analysis. Significance threshold was set at $\alpha < 0.05$.

Small RNA library construction

The same samples that were used in microarray study (14 IbxMe F_2 sows) were employed to construct small RNA libraries. Total RNA was extracted using TRIzol® reagent (Invitrogen) following the manufacturer's instructions. RNA quantity and integrity was assessed as previously described. Only total RNA with RIN $\geq$ 7 was used. Small RNA libraries were constructed by using 50 µg of total RNA from each sample. The small RNA fraction (15 - 30 nt) was separated on 12.5 % denatured polyacrylamide gels using miSpike™ internal control RNA (IDT®, Coralville, USA) as size internal marker. Small RNAs isolated were recovered from polyacrylamide gels and ligated to a 3' linker from miRCat™ kit (IDT®, Coralville, USA) using T4 RNA ligase without ATP (Fermentas, Germany). The ligated products (~ 45 nt) were eluted from 12.5 % denatured polyacrylamide gels to remove unligated products. Subsequently, the 3' linked RNAs were ligated with a 5' linker from miRCat™ kit (IDT®, Coralville, USA) using a T4 RNA ligase with ATP (Ambion, Austin, USA). The 5' and 3' linked RNAs products (~ 60 nt) were used to perform a reverse transcriptase reaction by using SuperScriptTM III Reverse Transcriptase kit (Invitrogen, Carlsbad, USA). Afterwards, cDNA were amplified by PCR with Platinum® Taq DNA Polimerase High Fidelity kit (Invitrogen, Carlsbad, USA). PCR products were purified by QIAquick® PCR Purification Kit (Quiagen®, Germany) and were quantified with Quant-IT™ using Qubit™ fluorometer (Invitrogen, Carlsbad, USA). Finally, each library was deeply sequenced by 454 Genome Sequencer FLX Titanium (Roche, CT, USA) available at DNA sequencing facilities at CRAG.

Annotation of sequences obtained by deeply sequencing

Primer sequences were trimmed and only insert sequences between 15 and 30 nucleotides with a total number of sequences 2.4 (taking all 14 libraries into

consideration) were left for further analysis. The remaining sequences were firstly compared to all available sequences in miRBase v18 (November 2011) using local Blast. Parameters were set to 100 % identity and up to 4 mismatches at the end of the sequences to assume variability on 3' ending. Sequences without homology with any miRBase v18 sequence were aligned to Rfam database (Rfam 10.1, June 2011) using local Blast. The blast parameters were set to up to 1 mismatch for miRNA sequences and up to 2 mismatches for other non-coding RNAs, and no more than 4 mismatches at the sequence ends in both cases. Finally, sequences without homology in Rfam database were compared to the complete GenBank database (NCBI-blast, March 2010) using local Blast. Hits were filtered accepting to up to 1 mismatch for miRNA sequences and up to 2 mismatches within the alignment for other RNAs, and up to 4 mismatches at the sequence ends.

miRNA RT-qPCR

For miRNA RT-qPCR analyses, total RNA from uterus was extracted by using the same protocol as used for small RNA libraries construction. The RNA quality and the quantification was performed as in microarray study. The expression of miRNAs was measured using the same sample size as used for microarray RT-qPCR validation (n = 36) but one sample from low prolificacy group was eliminated due to technical problems. Overall, a total of 35 IbxMe F_2 sows which include the 14 sows employed in the small RNA libraries construction. IbxMe F_2 sows were divided into two groups according the NE at 30-32 days of gestation; 16 high (NE ≥ 13) and 19 from low (NE ≤ 11) prolificacy sows. All samples used have a RIN ≥ 7.

Thirteen miRNAs were selected as candidate miRNAs to be evaluated through RT-qPCR. Five miRNAs (miR-125b-5p, miR-23b-3p, miR-200c-3p, miR-200b-3p and miR-23a-3p) were selected due to the fact that they were the most abundant miRNAs in uterus of pregnant sows; whereas the other eight (miR-139-5p, miR-150-5p, miR-27a-3p, miR-20-5p, miR-21-5p, miR101-3p, miR-122-5p and miR-199b-3p) were chosen because they have been reported to play important role in reproductive processes (Lei *et al.*, 2011,Abd El Naby *et al.*, 2011,Carletti and Christenson, 2009)

and were, therefore, considered biological candidate miRNAs. miR-16-5p, miR-103-3p, miR-25-3p and let-7a-5p were used as reference miRNAs based on the literature (Peltier and Latham, 2008).

Reverse transcription reactions were performed in duplicate with the universal cDNA synthesis kit (Exiqon, Denmark) using 100 ng of total RNA following the manufacter's instructions. qPCRs were performed using miRCURY^{TM} LNA^{TM} probes (Exiqon, Denmark) on an MX3000P machine (Stratagene, Le Jolla, CA, USA). Each qPCR was done in 10 µL total volume including 5 µl of QuantiFast SYBR Green PCR master mix (Qiagen®, Germany), 0.25 µL (except 0.5 µL for miR-125b-5p, miR-200c-3p, miR-150-5p, miR-21-5p, let-7a, miR-16, miR-103 and miR-25) of each primer spiked with LNA (Exiqon, Denmark) and 1 µl of a 1:20 dilution of the cDNA. Thermal profile was 95 ºC for 10 min followed by 40 amplification cycles of 95°C for 30 sec and 60°C for 60 sec. Standard curve with 10-fold dilution (made with equal amount of cDNA from all the samples) was included in all qPCR assays to calculate qPCR efficiencies. Melting curve analysis was included in each qPCR in order to avoid unspecific amplifications.

RT-qPCR Data Analysis

GeNorm algorithm (Vandesompele *et al.*, 2002) was used to evaluate the stability of the reference miRNAs (miR-16-5p, miR-103-3p, miR-25-3p and let-7a-5p) used ($M<1.5$) and to obtain a normalization factor (NF) based on the combined information of all reference miRNAs used. Data from RT-qPCR was analyzed using quantities obtained from quantification cycle (Cq) and using standard curve in order to correct differences in amplification efficiencies. Quantities for each sample were normalized by NF and fold changes were then calculated based on the lowest normalized quantity. Fold changes were Log_2 transformed and used for statistical analysis. Differential miRNA expression was performed using an Analysis of Variances (ANOVA) taking two factors into account: retrotranscription's replicate (RT-1 and RT-2) and prolificacy level (High and Low). Data analysis was performed using the PROC GLM of the Statistical Analysis System (Statistics, V 9.1.3; SAS Institute, Inc., Cary, NC). Significance threshold was set at $\alpha < 0.05$.

Gene Target Identification

Putative target mRNAs for prolificacy-related miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p) were identified with the *Ingenuity Pathway Analysis* (IPA) software (Ingenuity® Systems, USA).

RESULTS

Differential gene expression of high and low prolificacy sows by microarray analysis

The comparison of the gene expression level between high and low prolificacy sows in uterus samples revealed 351 differentially expressed (DE) probes (FDR < 0.05, Log_2 Fold Change > 2, Supplemental Table 1). Some probes targeted different regions of the same gene, showing similar patterns of expression. Overall, 113 probes (101 genes) were up-regulated and 238 probes (196 genes) were down-regulated in the high prolificacy group. Probes could be assigned to unique genes, and were therefore used for further *in silico* functional analyses. Gene ontology (GO) analysis was performed with the DAVID software and only significantly GO were considered. In high prolificacy sows, the most enriched GO categories for up-regulated genes were related to inflammatory response to stimulus and muscle tissue development processes, whereas for down-regulated genes the main GO enriched categories were related to muscle tissue development, cell junction organization, biological regulation, adhesion, muscle system process, circulatory system process and transport processes.

Validation of microarray experiments by quantitative PCR

To confirm results obtained in the microarray experiment, 4 genes (*BMP2, CALB3, PKP2* and *PTGIS*) were selected for RT-qPCR validation in the uterus dataset. These

genes were selected because of their localization on SSC13 where a QTL for litter size was reported in the same population and also by their role on the reproductive function. *BMP2*, *CALB3* and *PKP2* appeared to be differentially expressed based on the selected cut off (Supplementary Table 1); *PTGIS*, however, did not reach the Fold Change > 2 but it was very close (Fold Change = 1.9) and it had a significant adjusted *P-value* of 0.004 and was, therefore, selected for validation by RT-qPCR.

RT-qPCR results are shown in Table 2. All RT-qPCR validated genes displayed significant differences in gene expression between high and low prolificacy sows and showed good agreement in the expression direction with the microarray analysis. However, the fold changes determined with RT-qPCR were higher than those determined by microarray analysis, thus evidencing that RT-qPCR is a more accurate technique to measure the expression levels. Overall, validation of mRNA expression levels by RT-qPCR indicated that the microarray data was highly reliable although precise expression levels measures are obtained by RT-qPCR. Furthermore, in order to confirm these results, the sample size was expanded (n = 36; 20 from low prolificacy (NE≤ 11) and 16 from high prolificacy (NE≥13)). *BMP2* was still significantly differentially expressed (*P-Value* < 0.05) and *CALB3* and *PKP2* were suggestively differentially expressed (*P-Value* < 0.1) (Table 2). However, differential expression for *PTGIS* could not be determined (Table 2).

Annotation of small RNAs

In order to identify the miRNAs expression profile in uterus of pregnant sows regarding prolificacy levels, the same 14 IbxMe F_2 sows employed in microarray study were used for the construction of small RNA libraries and were deeply sequenced by 454 pyrosequencing (Roche). After removing the reads of low quality, trimming the adaptor sequence and eliminating the sequences with low copy number of sequences (CN≤ 3), a total of 172,901 sequences (2,035 unique sequences) from 15 to 30 nucleotides in length were obtained. From the size distribution of total reads, sequences between 21 - 23 nucleotides, the range length described for miRNAs, represented the 86.66 % of total reads. The sequences were

annotated into different categories by using different databases (miRBase v18, Rfam 10.0 and NCBI GenBank) (Table 3). By this way, 164,854 sequences (1,544 unique sequences) were annotated as miRNAs (95.3 %). The remaining 8,047 sequences (491 unique sequences, 4.7 %) were classified as other small non coding RNAs like small nucleolar RNA (snoRNA) (0.1 %), small nuclear RNA (snRNA) (0.1 %), coding RNA products resulting from RNA degradation such as tRNAs (1.4 %), rRNAs (0.5 %) and mRNA (1.6 %), RNA from bacteria and virus (0.1 %) and other RNAs (0.3 %) (Table 3). Interestingly, 999 sequences (93 unique sequences) could not be annotated and were suggested to be novel miRNAs candidates although further studies must be performed in order to validate if these sequences are novel miRNAs (Table 3). Overall, it is showed that the small RNA libraries were highly enriched in miRNAs. Thus, it is evidenced that the protocol used is highly efficient for miRNA detection.

miRNA expression profile

A total of 188 mature miRNAs (155 miRNAs) were identified in low prolificacy sows whereas 178 mature miRNAs (147 miRNAs) in high prolificacy sows (Table 4).

The most abundant uterine miRNA in both prolificacy groups was miR-125b-5p. Furthermore, miR-200c-3p, miR-200b-3p, miR-23b-3p, miR-23a-3p, miR-99a-5p were highly expressed although their relative abundance varied slightly depending on the prolificacy group (Table 4). The seventh miRNA most expressed in low prolificacy sows was miR-145-5p (eleventh in high prolificacy sows) whereas in high prolificacy sows was miR-126-5p (tenth in low prolificacy sows) (Table 4). Overall, the top 7 most abundant miRNAs represent the 70.51 % and the 67.44 % among total miRNAs for high and low prolificacy sows; respectively.

In general, most of the mature miRNAs (166 mature miRNAs) were expressed showing small Log Fold Changes (LogFC≤ 2) between high and low prolificacy sows (Figure 1). Only 13 mature miRNAs displayed higher LogFC between groups (37≥4 LogFC ≤ 2.1). However, these high LogFC were obtained when the LogConcentration

for the mature miRNAs was very low (Figure 1). In these cases, we believed that the high LogFC were overestimated and they may not be reliable.

miRNA RT-qPCR

Thirteen miRNAs were selected to be profiled by RT-qPCR. Five miRNAs (miR-125b-5p, miR-23b-3p, miR-200c-3p, miR-200b-3p and miR-23a-3p) were selected due to the fact that they were the most expressed in pregnant uterus. The other eight miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p, miR-20-5p, miR-21-5p, miR101-3p, miR-122-5p and miR-199b-3p), considered as biological candidate miRNAs, were chosen because they had been related to reproductive processes in the literature (Lei *et al.*, 2011,Abd El Naby *et al.*, 2011,Carletti and Christenson, 2009). The expression of miRNAs was measured in uterus of 35 IbxMe F_2 sows displaying extreme values for prolificacy levels: high (NE≥ 13; n = 16) and low (NE ≤ 11; n = 19) prolificacy sows which included the 14 IbxMe F_2 sows used for microarray and miRNA expression profile studies. miR-16-5p, miR-103-3p, miR-25-3p and let-7a-5p5p were used as reference miRNAs due to their expression stability (M < 1.5) assessed by Genorm software (Vandesompele *et al.*, 2002).

Expression data by RT-qPCR revealed similar expression levels between both prolificacy groups in the most expressed miRNAs (miR-125b-5p, miR-23b-3p and miR-23a-3p) (Figure 2) which is in concordance with the expression levels obtained by high throughput sequencing (Table 4). However, although data from high throughput sequencing suggest that miR-200c-3p, and miR-200b-3p could be differentially expressed regarding prolificacy levels (Table 4), the measure of their expression levels through RT-qPCR show that both miRNAs are expressed equally among both prolificacy groups (Figure 2). Among biological candidate miRNAs, in five out of the nine biological candidate miRNAs (miR-21-5p, miR-101-3p, miR-122-5p, let-7d-5p and miR-199b-3p) no differential expression was determined between prolificacy groups (Table 5). These results are in agreement with the high throughput sequencing data (Table 4). On the other hand, in four biological candidate miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p) expression data by RT-qPCR

showed differential expression regarding prolificacy levels (Table 5) even though high throughput data only suggest differential expression for miR-150-5p (Table 4). miR-139-5p and miR-150-5p were differentially expressed (*P-value* < 0.05) with 1.2 and 1.3 fold-increased in low compared with high prolificacy sows, respectively (Table 5). Expression of miR-27a-3p and miR-20-5p were suggestively greater in low prolificacy than in high prolificacy sows (*P-value* < 0.1) with 1.2 and 1.3 fold-increased (Table 5). All four miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p) were called prolificacy-related miRNAs.

mRNA target prediction of prolificacy-related miRNAs

mRNA targets for prolificacy-related miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p) were predicted *in silico* by the Ingenuity Pathways Analysis software (Ingenuity® Systems, USA). Overall, prolificacy-related miRNAs could target a total of 4,477 mRNAs; 1,883, 1,749, 1,119 and 841 for miR-20-5p, miR-27a-3p, miR-150-5p and miR-139-5p, respectively. Using expression mRNA data, 94 mRNA differentially expressed were determined as putative targets for prolificacy-related miRNAs (42, 34, 22 and 21 for miR-20-5p, miR-27a-3p, miR-139-5p and miR-150-5p) (Table 6). Among these genes, 69 down-regulated and 25 up-regulated genes in low prolificacy sows were predicted for the four upregulated prolificacy-related miRNAs. Together, the four prolificacy-related miRNAs could regulate the 31.64 % of the differentially expressed genes in uterus of pregnant sows regarding prolificacy levels. The expression level of the four prolificacy-related miRNAs was correlated with the expression levels of their DE targets. A total of 40 significant correlations were identified (*P-value* < 0.05). Interestingly, similar correlations were determined when different probes by one gene were analyzed. Surprisingly, the number of positive correlation was higher than the negative correlation; 28 positive *versus* 12 negative correlations (Table 6). Mainly, correlations involve miR-139-5p and miR-27a-3p.

DISCUSSION

The present study describes the transcriptome and the microRNAome in uterus of pregnant IbxMe F_2 sows displaying extreme phenotype for prolificacy levels (defined as the number of embryos attached to the uterus at day 30-32 of gestation). Moreover, an integrated overview of the mRNA and miRNA expression profile is provided.

To our knowledge, this is the first study that provides an overview of the transcriptome and the microRNAome in uterus of pregnant sows regarding prolificacy levels. Importantly, at day 31 of the gestation, the foetal survival rate will be an indication of the final litter size (Spotter and Distl, 2006). Thus, the NE measured at day 30-32 of gestation could be taken as an estimation of prolificacy. Differential expression in 297 genes was determined between high and low prolificacy sows. *In silico* functional analyses in high prolificacy sows suggested that up-regulated genes are mainly involved in inflammatory response to stimulus and in muscle tissue development processes whereas down-regulated genes are related to cell junction organization and adhesion, biological regulation, muscle and circulatory system process, transport and in muscle tissue development. Porcine conceptuses are highly invasive with high enzymatic potential to erode into the uterine tissue (Geisert and Yelich, 1997). In order to protect uterine epithelial cells from proteolytic conceptuses, a thickening of the uterine epithelial glycocalyx is created (Keys and King, 1990,Ferrell *et al.*, 2003). Thus, in high prolificacy sows, the up-regulation of genes involved in muscle tissue development might participate in the creation of the barrier between the proteolytic trophoblast and uterine epithelial cells in order to protect the uterine surface epithelium which would benefit the embryo survival. In this protective barrier created, it has been reported that the presence of proteins like mucins with anti-adhesive properties are very important to protect the endometrium from microbial invasion (Carraway *et al.*, 2002) as well as from proteolytic activity of porcine conceptus (Ferrell *et al.*, 2003). The important presence of anti-adhesive proteins would be in agreement with down-regulation of genes associated with cell junction organization and adhesion processes determined in high prolificacy sows in order to create a better uterine environment to benefit

the embryo development and, therefore, the embryo survival. These endometrial responses evoked by the developing pig conceptus during uterine attachment resemble the acute phase response induced during generalized tissue inflammation (Geisert and Yelich, 1997,Salier *et al.*, 1996) that would explain the up-regulation of genes related to the inflammatory response in high prolificacy sows. On the other hand, down-regulated genes in high prolificacy sows were associated with muscle and circulatory system process, protein transport and muscle tissue development. This result was not expected due to the fact that these processes have been described as beneficial for embryo development (Croy *et al.*, 2009,Tayade *et al.*, 2007). However, overexpression of these genes could lead the embryo loss and, thus, their expression might be tightly regulated. Additionally, it could be that these genes have functions that have not been described yet which would be important for embryo survival. Further work must be performed in order to decipher if these genes have other functions.

miRNAs have been described as post-transcriptional gene expression regulators with important functions in a wide range of biological processes including reproduction (Nothnick, 2012,Hawkins *et al.*, 2011,Carletti and Christenson, 2009) although their role in porcine pregnancy regarding prolificacy levels is still little. The miRNA expression profile determined that the 6 top most abundant miRNAs in both prolificacy groups were miR-125b-5p, miR-23b-3p, miR-200c-3p, miR-200b-3p, miR-23a-3p and miR-99a-5p. According to our results, miR-125b, miR-23b and miR-99a have also been described as the most abundant miRNAs in female uterus (Pan *et al.*, 2007) and miR-200 family members are significantly induced in uterus during late gestation in humans and mouse where they mediate uterine contraction (Renthal *et al.*, 2010). miR-23a as well as miR-23b have been previously described in human uterine cells where they can be regulated by steroids hormones (Toloubeydokhti *et al.*, 2008) which are of main importance in the regulation of reproductive processes. However, miR-23a-3p is the first time that its high abundance in porcine pregnant uterus is reported. It has been suggested that uterine miRNAs are related to processes that include inflammation, apoptosis, proliferation, angiogenesis and differentiation (Nothnick, 2012) which are in agreement with the function associated

with the most expressed miRNAs in uterus (Nothnick, 2012,Ruan *et al.*, 2012,Jiang *et al.*, 2011,Lerman *et al.*, 2011,Oneyama *et al.*, 2011,Zhou *et al.*, 2011,Wang *et al.*, 2010). Moreover, the dysregulation of them has been associated with uterus pathologies like endometriosis and endometrial cancers (Nothnick, 2012,Pan *et al.*, 2007,Jiang *et al.*, 2011,Park *et al.*, 2012,Lee *et al.*, 2011,Ohlsson Teague *et al.*, 2009).

miRNA expression profile was determined using a high throughput sequencing methodology. Although it is known that this technique could be used for measure gene expression, the most important concerns about sequencing miRNA is the depth of sequencing required to effectively sample the miRNAome and to perform a reliable miRNA differential expression analysis (Malone and Oliver, 2011). In a previous miRNA study performed by our group in ovary of pregnant sows, we could determine that differences in miRNAs expression regarding prolificacy levels are very low. Thus, we believed that the number of total sequences reads obtained in the present work was insufficient to perform a good differential expression study and, therefore, it can only be used to describe the miRNAs expressed in uterus. However, through the literature, some miRNAs have been described as important for reproductive processes (i.e. miR-139-5p, miR-150-5p, miR-27a-3p, miR-20-5p, miR-21-5p, miR101-3p, miR-122-5p, let-7d-5p and miR-199b-3p) (Lei *et al.*, 2011,Abd El Naby *et al.*, 2011,Carletti and Christenson, 2009). Among these, high throughput sequencing data only suggested differential expression for miR-150-5p. However, the expression levels of these eight miRNA were measured through RT-qPCR. Differential expression regarding prolificacy levels were detected for miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p and were, therefore, called prolificacy-related miRNAs. Small differences in fold change (from 1.2 to 1.3) between high and low prolificacy were detected. It is in concordance with other miRNA differential expression studies related to reproductive processes where the differences in miRNAs fold changes are small, ranging from 1.5 to 2.5 (Su *et al.*, 2010b, Balcells *et al.*, 2012). Although differences in fold changes are small, they may be biologically significant due to the fact that miRNAs regulate many different mRNA target and thus, the effect is amplified and may result in large phenotypic differences (Calin and Croce, 2006). As it is mentioned above, all four prolificacy-related miRNAs have been

shown to play key roles in reproductive processes. miR-150-5p is involved in bovine oocyte maduration and in early embryo development (Abd El Naby *et al.*, 2011). In pigs, miR-20a-5p and miR-27a-3p are differentially expressed in placenta between day 30 and 90 of gestation; miR-20-5p is upregulated whereas miR-27a-3p is downregulated at 30 days of gestation. These miRNAs might be important for placental growth and function (Su *et al.*, 2010a). Furthermore, miR-27a-3p is differentially expressed in ovaries between Large White and Chinese Erhualian breeds and interestingly, a SNP described at 18bp downstream of pre-miRNAs has been associated with litter size in DIV pigs population (4th Dam lines of Chinese lean-type new lines) (Lei *et al.*, 2011). Aberrant expression of prolificacy-related miRNAs has been related to some reproductive diseases which cause infertility. miR-139-5p and miR-20a have been associated with placenta preeclampsia (Wang *et al.*, 2012,Enquobahrie *et al.*, 2011), miR-150-5p and miR-20-5p with endometriosis disease (Ohlsson Teague *et al.*, 2009) and miR-139b-5p with leiomyomas (Marsh *et al.*, 2008). On the other hand, it has been described that prolificacy-related miRNAs can control reproductive functions by regulating ovarian steroids hormones (Sirotkin *et al.*, 2009).

The identification of the targets for miRNAs is a great challenge in order to better understand the biological role of miRNAs. For prolificacy-related miRNAs, a total of 4,477 putative mRNAs targets were determined. However, we focused on the candidate mRNAs targets that were described as differentially expressed in order to understand if the dysregulation of these genes could be due to prolificacy-related miRNAs and thus, determine new molecular mechanisms related to prolificacy levels. Among the differentially expressed genes, 96 genes could be target for prolificacy-related miRNAs. In general, it is believed that miRNAs bind to the 3' untranslated region of target mRNAs and down-regulate their expression through mRNA degradation and/or by repressing the mRNA translation (Huntzinger and Izaurralde, 2011). The significant negative correlation found between miR-27a-3p and miR-139-5p suggests that these miRNAs have a direct effect on their targets by modulating their mRNA expressions levels. Therefore, the down-regulation of these miRNAs would explain the up-regulation of their targets in high prolificacy sows.

However, the major significant correlations detected between miR-27a-3p and miR-139-5p and their differentially expressed targets were positive. In these cases, miRNAs might regulate their targets positively. Although it seems in disagreement with the general role suggested for miRNAs, it has been reported that some miRNAs can also be positive regulators of their targets (Vasudevan, 2011,Vasudevan *et al.*, 2007). On the other hand, these positive correlations could be explained by the fine tunning role associated with miRNAs that decrease the expression level of their targets slightly over time and at time that we measure the expression levels, the down-regulation of their target was not reflected.

Prolificacy-related miRNAs, miR-139-5p, miR-150-5p, miR-27a-3p, miR-20-5p, are located on SSC9, SSC6, SSC11 and SSC12, respectively (Pre Ensembl, Sscrofa10.2). Although miR-139-5p and miR-150-5p are located in chromosomes where epistatic QTL interactions for prolificacy-related traits have been reported (Noguera *et al.*, 2009), they are not within the confidence intervals. However, miRNAs are regulatory molecules and they could regulated genes which could be candidates for the prolificacy QTL described (Noguera *et al.*, 2009,Fernandez-Rodriguez *et al.*, 2010). In this sense, the 94 differentially expressed genes that are putative targets for prolificacy-related miRNAs were mapped in porcine genome (Sscrofa 10.2) in order to identify potential candidate genes for prolificacy-related miRNAs. A total of 32 genes were located within confidence intervals of significant QTL affecting the number of piglets born alive (NBA) and the total number of piglets born (TNB) (Table 7). Among these, 10 genes (*AKAP2*, *ITGA5*, *PTHLH*, *DSP*, *CAP2*, *HLA-B*, *NR2F2, IL1RAP, MUC4* and *BMP2* genes) have been previously associated with reproductive processes (Balcells *et al.*, 2011,Balcells *et al.*, 2011,Balcells *et al.*, 2011,Merkl *et al.*, 2010,Tajiri *et al.*, 2010,Le Bouteiller *et al.*, 2009,Ross *et al.*, 2009,Singh *et al.*, 2008,Jasoni *et al.*, 2005,Dziechciowski and Klimek, 2001,Moss and Gerton, 2001,Illingworth *et al.*, 2000). However, there are candidate genes which have never been associated with reproductive function (*NCS1*, *CALD1, MYO5B*, *GEM*, *PKP2*, *CFL2*, *RECQL5*, *PALLD*, *MGLL*, *MYLK*, *JPH2* and *FADS6* genes) or even their function is unknown (*RNGTT*, *LRP1B, PKIA*, *AEBP2*, *NBL1*, *CAMK2N1*, *ADAMTSL3*, *NME7*, *SYNPO2* and *SCHIP1* genes). During pregnancy, it has been suggested that adhesion

molecules, cytokines and lipids are involved in the early fetal-maternal interactions (Vigano *et al.*, 2003) which are in fact, the product of the candidate genes that have been related with reproductive processes. Interestingly, candidate genes without a previously described relation to reproductive processes have the same functions as those described for genes which have been associated with reproductive function. Thus, they are good candidate to evaluate their effect on the prolificacy-related traits. Importantly, miR-150-5p has been described as a novel regulator of *MUC4* gene (Srivastava *et al.*, 2011) which is located within the confidence interval for prolificacy QTL on SSC13 and it has been associated with litter size (Balcells *et al.*, 2011). In high prolificacy sows, the up-regulation of *MUC4* might establish an optimal uterine environment essential for adequate embryo development (Balcells *et al.*, 2011) whereas miR-150-5p is downregulated. It is hypothesized that miR-150-5p could regulate by a novel post-transcriptional mechanism the expression level of *MUC4* gene. However, no significant negative correlation was determined between miRNAs and *MUC4* gene expression but it could be that miRNA acts as a fine-tunning mechanism and the decrease level of mRNA expression could not be detected at time of measure. However, further studies must be performed in order to confirm these results.

CONCLUSIONS

Differential expression was determined at transcriptome and microRNAome level regarding prolificacy levels suggesting an important role of all of them in reproductive processes.

ACKNOWLEDGEMENTS

The authors are indebted to Luis Varona, M. Arqué, J. Tarrés, M. Fina, and the staff of Nova Genètica, in particular to E. Ramells, F. Márquez, R. Malé, F. Rovira, and I. Riart, for cooperating in the experimental protocol. The authors gratefully acknowledge INRA (France) and the CIA El Dehesón del Encinar (Spain) for providing

the purebred Meishan sows and Iberian boars, respectively. This research was funded in part by Project AGL2004-08368-C03, AGL2007-66371-C02-01, AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 Program (CSD2007-00036), both from the Spanish Ministry of Science and Innovation. IB is recipient of PIF PhD fellowship from Universitat Autònoma de Barcelona.

REFERENCES

Abd El Naby W.S., Hagos T.H., Hossain M.M., Salilew-Wondim D., Gad A.Y., Rings F., Cinar M.U., Tholen E., Looft C., Schellander K., Hoelker M. & Tesfaye D. (2011) Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. Zygote (Cambridge, England), 1-21.

Alminana C., Heath P.R., Wilkinson S., Sanchez-Osorio J., Cuello C., Parrilla I., Gil M.A., Vazquez J.L., Vazquez J.M., Roca J., Martinez E.A. & Fazeli A. (2012) Early developing pig embryos mediate their own environment in the maternal tract. PloS one 7, e33625.

Balcells I., Castello A., Mercade A., Noguera J.L., Fernandez-Rodriguez A., Sanchez A. & Tomas A. (2011) Analysis of porcine MUC4 gene as a candidate gene for prolificacy QTL on SSC13 in an Iberian x Meishan F2 population. BMC genetics 12, 93.

Balcells I., Timoneda O., Cirera S., Søkilde R., Litman T., Fernández-Rodríguez A., Noguera, J.L., Sánchez A. & Tomás A. (2012) Differential microRNA expression in ovary of high and low prolificacy pregnant sows. Submitted.

BENJAMINI Y. & HOCHBERG Y. (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing Rid C-4219-2008. Journal of the Royal Statistical Society Series B-Methodological 57, 289-300.

Blomberg le A., Schreier L. & Li R.W. (2010) Characteristics of peri-implantation porcine concepti population and maternal milieu influence the transcriptome profile. Molecular reproduction and development 77, 978-989.

Caetano A.R., Johnson R.K., Ford J.J. & Pomp D. (2004) Microarray profiling for differential gene expression in ovaries and ovarian follicles of pigs selected for increased ovulation rate. Genetics 168, 1529-1537.

Calin G.A. & Croce C.M. (2006) MicroRNA-cancer connection: the beginning of a new tale. Cancer research 66, 7390-7394.

Carletti M.Z. & Christenson L.K. (2009) MicroRNA in the ovary and female reproductive tract. Journal of animal science 87, E29-38.

Carraway K.L., Perez A., Idris N., Jepson S., Arango M., Komatsu M., Haq B., Price-Schiavi S.A., Zhang J. & Carraway C.A. (2002) Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. Progress in nucleic acid research and molecular biology 71, 149-185.

Croy B.A., Wessels J.M., Linton N.F., van den Heuvel M., Edwards A.K. & Tayade C. (2009) Cellular and molecular events in early and mid gestation porcine implantation sites: a review. Society of Reproduction and Fertility supplement 66, 233-244.

Dziechciowski M. & Klimek R. (2001) Comparison of periovulatory and early pregnancy blood levels of oxytocinase (CAP1) and isooxytocinase (CAP2). Early pregnancy (Online) 5, 113-120.

Enquobahrie D.A., Abetew D.F., Sorensen T.K., Willoughby D., Chidambaram K. & Williams M.A. (2011) Placental microRNA expression in pregnancies complicated by preeclampsia. American Journal of Obstetrics and Gynecology 204, 178.e12-178.e21.

Esteve-Codina A., Kofler R., Palmieri N., Bussotti G., Notredame C. & Perez-Enciso M. (2011) Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. BMC genomics 12, 552.

Fernandes L.T., Tomas A., Bensaid A., Perez-Enciso M., Sibila M., Sanchez A. & Segales J. (2009) Exploratory study on the transcriptional profile of pigs subclinically infected with porcine circovirus type 2. Animal Biotechnology 20, 96-109.

Fernandez-Rodriguez A., Munoz 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. Biology of reproduction 84, 299-307.

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. Animal Genetics 41, 73-80.

Ferrell A.D., Malayer J.R., Carraway K.L. & Geisert R.D. (2003) Sialomucin complex (Muc4) expression in porcine endometrium during the oestrous cycle and early pregnancy. Reproduction in domestic animals = Zuchthygiene 38, 63-65.

Geisert R.D. & Yelich J.V. (1997) Regulation of conceptus development and attachment in pigs. Journal of reproduction and fertility.Supplement 52, 133-149.

Gentleman R.C., Carey V.J., Bates D.M., Bolstad B., Dettling M., Dudoit S., Ellis B., Gautier L., Ge Y., Gentry J., Hornik K., Hothorn T., Huber W., Iacus S., Irizarry R., Leisch F., Li C., Maechler M., Rossini A.J., Sawitzki G., Smith C., Smyth G., Tierney L., Yang J.Y. & Zhang J. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome biology 5, R80.

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. Journal of animal science 82, 17-31.

Hawkins S.M., Buchold G.M. & Matzuk M.M. (2011) Minireview: The Roles of Small RNA Pathways in Reproductive Medicine. Molecular endocrinology (Baltimore, Md.).

Huntzinger E. & Izaurralde E. (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature reviews.Genetics 12, 99-110.

Illingworth I.M., Kiszka I., Bagley S., Ireland G.W., Garrod D.R. & Kimber S.J. (2000) Desmosomes are reduced in the mouse uterine luminal epithelium during the preimplantation period of pregnancy: a mechanism for facilitation of implantation. Biology of reproduction 63, 1764-1773.

Irizarry R.A., Hobbs B., Collin F., Beazer-Barclay Y.D., Antonellis K.J., Scherf U. & Speed T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) 4, 249-264.

Ishibashi O., Ohkuchi A., Ali M.M., Kurashina R., Luo S.S., Ishikawa T., Takizawa T., Hirashima C., Takahashi K., Migita M., Ishikawa G., Yoneyama K., Asakura H., Izumi A., Matsubara S., Takeshita T. & Takizawa T. (2012) Hydroxysteroid (17-beta) dehydrogenase 1 is dysregulated by miR-210 and miR-518c that are aberrantly expressed in preeclamptic placentas: a novel marker for predicting preeclampsia. Hypertension 59, 265-273.

Medhora M., Bousamra M., Zhu D., Somberg L. & Jacobs E.R. (2002) Upregulation of collagens detected by gene array in a model of flow-induced pulmonary vascular remodeling. Am J Physiol Heart Circ Physiol 282, 414-422.

Jasoni C.L., Todman M.G., Han S.K. & Herbison A.E. (2005) Expression of mRNAs encoding receptors that mediate stress signals in gonadotropin-releasing hormone neurons of the mouse. Neuroendocrinology 82, 320-328.

Jiang F., Liu T., He Y., Yan Q., Chen X., Wang H. & Wan X. (2011) MiR-125b promotes proliferation and migration of type II endometrial carcinoma cells through targeting TP53INP1 tumor suppressor in vitro and in vivo. BMC cancer 11, 425.

Keys J.L. & King G.J. (1990) Microscopic examination of porcine conceptus-maternal interface between days 10 and 19 of pregnancy. The American Journal of Anatomy 188, 221-238.

Le Bouteiller P., El Costa H., Aguerre-Girr M. & Tabiasco J. (2009) Immunity of pregnancy: novel concepts. Bulletin de l'Academie nationale de medecine 193, 1029-41; discussion 1041-2, 1067-8.

Lee J.W., Park Y.A., Choi J.J., Lee Y.Y., Kim C.J., Choi C., Kim T.J., Lee N.W., Kim B.G. & Bae D.S. (2011) The expression of the miRNA-200 family in endometrial endometrioid carcinoma. Gynecologic oncology 120, 56-62.

Lei B., Gao S., Luo L.F., Xia X.Y., Jiang S.W., Deng C.Y., Xiong Y.Z. & Li F.E. (2011) A SNP in the miR-27a gene is associated with litter size in pigs. Molecular biology reports 38, 3725-3729.

Lerman G., Avivi C., Mardoukh C., Barzilai A., Tessone A., Gradus B., Pavlotsky F., Barshack I., Polak-Charcon S., Orenstein A., Hornstein E., Sidi Y. & Avni D. (2011) MiRNA expression in psoriatic skin: reciprocal regulation of hsa-miR-99a and IGF-1R. PloS one 6, e20916.

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) 25, 402-408.

Malone J.H. & Oliver B. (2011) Microarrays, deep sequencing and the true measure of the transcriptome. BMC biology 9, 34.

Mardis E.R. (2008) Next-generation DNA sequencing methods. Annual review of genomics and human genetics 9, 387-402.

Marsh E.E., Lin Z., Yin P., Milad M., Chakravarti D. & Bulun S.E. (2008) Differential expression of microRNA species in human uterine leiomyoma versus normal myometrium. Fertility and sterility 89, 1771-1776.

Merkl M., Ulbrich S.E., Otzdorff C., Herbach N., Wanke R., Wolf E., Handler J. & Bauersachs S. (2010) Microarray analysis of equine endometrium at days 8 and 12 of pregnancy. Biology of reproduction 83, 874-886.

Moody D.E, Zou Z. & McIntyre L. (2002) Cross-species hybridisation of pig RNA to human nylon microarrays. BMC Genomics 27, 27.

Moss S.B. & Gerton G.L. (2001) A-kinase anchor proteins in endocrine systems and reproduction. Trends in endocrinology and metabolism: TEM 12, 434-440.

Noguera J.L., Rodriguez C., Varona L., Tomas A., Munoz 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.

Nothnick W. (2012) The role of microRNAs in the female reproductive tract. Reproduction (Cambridge, England).

Ohlsson Teague E.M., Van der Hoek K.H., Van der Hoek M.B., Perry N., Wagaarachchi P., Robertson S.A., Print C.G. & Hull L.M. (2009) MicroRNA-regulated pathways associated with endometriosis. Molecular endocrinology (Baltimore, Md.) 23, 265-275.

Oneyama C., Ikeda J., Okuzaki D., Suzuki K., Kanou T., Shintani Y., Morii E., Okumura M., Aozasa K. & Okada M. (2011) MicroRNA-mediated downregulation of mTOR/FGFR3 controls tumor growth induced by Src-related oncogenic pathways. Oncogene 30, 3489-3501.

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. Society of Reproduction and Fertility supplement 66, 87-102.

Ostrup E., Bauersachs S., Blum H., Wolf E. & Hyttel P. (2010) Differential endometrial gene expression in pregnant and nonpregnant sows. Biology of reproduction 83, 277-285.

Pan Q., Luo X., Toloubeydokhti T. & Chegini N. (2007) The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. Molecular human reproduction 13, 797-806.

Park Y.A., Lee J.W., Choi J.J., Jeon H.K., Cho Y., Choi C., Kim T.J., Lee N.W., Kim B.G. & Bae D.S. (2012) The interactions between MicroRNA-200c and BRD7 in endometrial carcinoma. Gynecologic oncology 124, 125-133.

Peltier H.J. & Latham G.J. (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. RNA (New York, N.Y.) 14, 844-852.

Renthal N.E., Chen C.C., Williams K.C., Gerard R.D., Prange-Kiel J. & Mendelson C.R. (2010) miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor. Proceedings of the National Academy of Sciences of the United States of America 107, 20828-20833.

Ross J.W., Ashworth M.D., Stein D.R., Couture O.P., Tuggle C.K. & Geisert R.D. (2009) Identification of differential gene expression during porcine conceptus rapid trophoblastic elongation and attachment to uterine luminal epithelium. Physiological genomics 36, 140-148.

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, 4761-4773.

Rothschild M.F., Hu Z.L. & Jiang Z. (2007) Advances in QTL mapping in pigs. International journal of biological sciences 3, 192-197.

Ruan W., Xu J.M., Li S.B., Yuan L.Q. & Dai R.P. (2012) Effects of down-regulation of microRNA-23a on TNF-alpha-induced endothelial cell apoptosis through caspase-dependent pathways. Cardiovascular research 93, 623-632.

Salier J.P., Rouet P., Raguenez G. & Daveau M. (1996) The inter-alpha-inhibitor family: from structure to regulation. The Biochemical journal 315 (Pt 1), 1-9.

Seemann S.E., Gilchrist M.J., Hofacker I.L., Stadler P.F. & Gorodkin J. (2007) Detection of RNA structures in porcine EST data and related mammals. BMC genomics 8, 316.

Schena M., Shalon D., Davis R.W. & Brown P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 20, 467-470.

Singh A.P., Castranio T., Scott G., Guo D., Harris M.A., Ray M., Harris S.E. & Mishina Y. (2008) Influences of reduced expression of maternal bone morphogenetic protein 2 on mouse embryonic development. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation 2, 134-141.

Sirotkin A.V., Ovcharenko D., Grossmann R., Lauková M. & Mlyn?ek M. (2009) Identification of MicroRNAs controlling human ovarian cell steroidogenesis via a genome-scale screen. Journal of cellular physiology 219, 415-420.

Smyth G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology 3, Article3.

Spotter A. & Distl O. (2006) Genetic approaches to the improvement of fertility traits in the pig. Veterinary journal (London, England : 1997) 172, 234-247.

Srivastava S.K., Bhardwaj A., Singh S., Arora S., Wang B., Grizzle W.E. & Singh A.P. (2011) MicroRNA-150 directly targets MUC4 and suppresses growth and malignant behavior of pancreatic cancer cells. Carcinogenesis 32, 1832-1839.

Su L., Zhao S., Zhu M. & Yu M. (2010a) Differential expression of microRNAs in porcine placentas on Days 30 and 90 of gestation. Reproduction, Fertility and Development 22, 1175-1182.

Su R.W., Lei W., Liu J.L., Zhang Z.R., Jia B., Feng X.H., Ren G., Hu S.J. & Yang Z.M. (2010b) The integrative analysis of microRNA and mRNA expression in mouse uterus under delayed implantation and activation. PloS one 5, e15513.

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.

Tajiri Y., Igarashi T., Li D., Mukai K., Suematsu M., Fukui E., Yoshizawa M. & Matsumoto H. (2010) Tubulointerstitial nephritis antigen-like 1 is expressed in the uterus and binds with integrins in decidualized endometrium during postimplantation in mice. Biology of reproduction 82, 263-270.

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. The Journal of reproduction and development 53, 455-463.

Toloubeydokhti T., Pan Q., Luo X., Bukulmez O. & Chegini N. (2008) The expression and ovarian steroid regulation of endometrial micro-RNAs. Reproductive sciences (Thousand Oaks, Calif.) 15, 993-1001.

Tsai S., Mir B., Martin A.C., Estrada J.L., Bischoff S.R., Hsieh W.P., Cassady J.P., Freking B.A., Nonneman D.J., Rohrer G.A. & Piedrahita J.A. (2006) Detection of transcriptional difference of porcine imprinted genes using different microarray platforms. BMC genomics 7, 328.

Tuggle C.K., Wang Y. & Couture O. (2007) Advances in swine transcriptomics. International journal of biological sciences 3, 132-152.

Uenishi H., Eguchi-Ogawa T., Shinkai H., Okumura N., Suzuki K., Toki D., Hamasima N. & Awata T. (2007) PEDE (Pig EST Data Explorer) has been expanded into Pig Expression Data Explorer, including 10 147 porcine full-length cDNA sequences. Nucleic acids research 35, D650-3.

Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. & Speleman F. (2002) Accurate normalization of real-time quantitative RT-PCR data by

geometric averaging of multiple internal control genes. Genome biology 3, RESEARCH0034.

Vasudevan S., Tong Y. & Steitz J.A. (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science (New York, N.Y.) 318, 1931-1934.

Vasudevan S. (2011) Posttranscriptional Upregulation by MicroRNAs. Wiley Interdisciplinary Reviews: RNA, n/a-n/a.

Velculescu V.E., Zhang L., Vogelstein B. & Kinzler K.W. (1995) Serial analysis of gene expression. Science 20, 484-487.

Vigano P., Mangioni S., Pompei F. & Chiodo I. (2003) Maternal-conceptus cross talk--a review. Placenta 24 Suppl B, S56-61.

Wang K.C., Garmire L.X., Young A., Nguyen P., Trinh A., Subramaniam S., Wang N., Shyy J.Y., Li Y.S. & Chien S. (2010) Role of microRNA-23b in flow-regulation of Rb phosphorylation and endothelial cell growth. Proceedings of the National Academy of Sciences of the United States of America 107, 3234-3239.

Wang W., Feng L., Zhang H., Hachy S., Satohisa S., Laurent L.C., Parast M., Zheng J. & Chen D.B. (2012) Preeclampsia Up-Regulates Angiogenesis-Associated MicroRNA (i.e., miR-17, -20a, and -20b) That Target Ephrin-B2 and EPHB4 in Human Placenta. The Journal of clinical endocrinology and metabolism.

Wilson C.L. & Miller C.J. (2005) Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics (Oxford, England) 21, 3683-3685.

Xu S., Linher-Melville K., Yang B.B., Wu D. & Li J. (2011) Micro-RNA378 (miR-378) regulates ovarian estradiol production by targeting aromatase. Endocrinology 152, 3941-3951.

Zhou Q., Gallagher R., Ufret-Vincenty R., Li X., Olson E.N. & Wang S. (2011) Regulation of angiogenesis and choroidal neovascularization by members of microRNA-23~27~24 clusters. Proceedings of the National Academy of Sciences of the United States of America 108, 8287-8292.

TABLES

Table 1. List of primer sequences used for validating microarray results by RT-qPCR.

Gene	Sequence (5'- 3')	Length (bp)
BMP2	Fw: ACGGTCTGCGGTCTCCTAAA Rv: GAAGCAGCAACGCTAGAAGACA	68
CALB3	Fw: GGAGGAGCTGAAGCAACTGATT Rv: CATCTAGGGTTCTCGGACCTTTC	68
PTGIS	Fw: GACCCGGCCATCTACACAGA Rv: CATCCGGGTTCAGGAATCG	61
PKP2	Fw: CGCACACAGAGCTCCATAATG Rv: CGGCTGTTGACAAAATCTGTCTT	67

Table 2. RT-qPCR expression level in uterus of genes selected based on microarray study. Gene expression levels were measured in pregnant IbxMe F_2 sows regarding prolificacy levels on day 30-32 of gestation.

Gene	Microarray validation (n = 14)		Further validation (n = 36)	
	Fold Change[a] High/Low	*P-Value*[b]	Fold Change[a] High/Low	*P-Value*[b]
BMP2	3.74	0.0042	1.97	0.0467
CALB3	8.25	0.0319	2.42	0.0821
PKP2	3.38	0.0004	1.44	0.1337
PTGIS	- 3.81	0.0032	- 1.16	0.6429

[a]Positive and negative signs indicate whether the level of gene expression is higher in the first or in the second group, respectively.

[b]Nominal *P-Value*. Significance threshold was set at 0.05.

Table 3. Annotation of the total number of small RNAs sequences from uterus libraries. Libraries were constructed from pregnant 14 IbxMe F_2 sows divided into two groups: high and low prolificacy sows according the number of embryos (NE) attached to the uterus on day 30-32 of gestation.

Annotation	Low prolificacy sows (n = 9; NE ≤ 11)		High prolificacy Sows (n = 5; NE ≥ 14)		Total	
microRNAs	113288	95.8%	51566	94.5%	164854	95.3%
Transfer RNAs	1306	1.1%	1097	2.0%	2403	1.4%
Ribosomal RNAs	571	0.5%	305	0.6%	876	0.5%
Small nuclear RNAs	111	0.1%	51	0.1%	162	0.1%
Small nucleolar RNAs	86	0.1%	44	0.1%	130	0.1%
Bacterial and viral RNAs	104	0.1%	48	0.1%	152	0.1%
Messenger RNAs	1889	1.6%	939	1.7%	2828	1.6%
Other RNAs[a]	307	0.3%	190	0.3%	497	0.3%
Unknown	647	0.5%	352	0.6%	999	0.6%
Total number of sequences	118309		54592		172901	

[a] Includes Y RNAs and Metazoa signal recognition particle RNA (SRP)

Table 4. miRNA expression profile in uterus of pregnant sows. miRNA expression profile was evaluated in 14 IbxMe F_2 regarding prolificacy levels according the NE on day 30-32 of gestation (high (n = 5; NE ≥ 14) and low (n = 9; NE ≤ 11)).

miRNA	miRNA Sequence[a]	Total		Low Prolificacy Sow (n = 9; NE ≤ 11)		High Prolificacy Sow (n = 5; NE ≥ 14)	
		CN[b]	%	CN[b]	%	CN[b]	%
let-7a-5p	TGAGGTAGTAGGTTGTATAGTT	1188	0.69%	886	0.75%	302	0.55%
let-7b-3p	CTATACAACCTACTGCCTTCCT	40	0.02%	30	0.03%	10	0.02%
let-7b-5p	TGAGGTAGTAGGTTGTGTGGTT	973	0.56%	721	0.61%	252	0.46%
let-7c-5p	TGAGGTAGTAGGTTGTATGGTT	238	0.14%	182	0.15%	56	0.10%
let-7d-3p	CTATACGACCTGCTGCCTTTCT	128	0.07%	71	0.06%	57	0.10%
let-7d-5p	AGAGGTAGTAGGTTGCATAGTT	284	0.16%	231	0.20%	53	0.10%
let-7e-5p	TGAGGTAGGAGGTTGTATAGTT	59	0.03%	43	0.04%	16	0.03%
let-7f-5p	TGAGGTAGTAGATTGTATAGTT	70	0.04%	53	0.04%	17	0.03%
let-7g-5p	TGAGGTAGTAGTTTGTACAGTT	5	0.00%	1	0.00%	4	0.01%
let-7i-5p	TGAGGTAGTAGTTTGTGCTGTT	69	0.04%	52	0.04%	17	0.03%
miR-1-3p	TGGAATGTAAAGAAGTATGTAT	12	0.01%	6	0.01%	6	0.01%
miR-9-3p	TAAAGCTAGATAACCGAAAGT	6	0.00%	6	0.01%	0	0.00%
miR-10a-5p	TACCCTGTAGATCCGAATTTGT	1739	1.01%	1134	0.96%	605	1.11%
miR-10b-5p	TACCCTGTAGAACCGAATTTGT	3889	2.25%	2855	2.41%	1034	1.89%
miR-10c-5p	TACCCTGTAGAATCGAATTTGT	5	0.00%	5	0.00%	0	0.00%
miR-15b-5p	TAGCAGCACATCATGGTTTAC	30	0.02%	21	0.02%	9	0.02%
miR-16-5p	TAGCAGCACGTAAATATTGGC	25	0.01%	18	0.02%	7	0.01%
miR-18a-5p	TAAGGTGCATCTAGTGCAGAT	24	0.01%	10	0.01%	14	0.03%
miR-19b-3p	TGTGCAAATCCATGCAAAACTGA	14	0.01%	11	0.01%	3	0.01%
miR-20-5p	TAAAGTGCTTATAGTGCAGGTAG	83	0.05%	62	0.05%	21	0.04%
miR-20b-5p	CAAAGTGCTCACAGTGCAGGTAG	4	0.00%	4	0.00%	0	0.00%
miR-21-5p	TAGCTTATCAGACTGATGTTGAC	1366	0.79%	1033	0.87%	333	0.61%
miR-22-3p	AAGCTGCCAGTTGAAGAACTGT	85	0.05%	53	0.04%	32	0.06%
miR-23a-3p	ATCACATTGCCAGGGATTTCCA	16083	9.30%	10443	8.83%	5640	10.33%
miR-23b-3p	ATCACATTGCCAGGGATTACCA	22253	12.87%	15698	13.27%	6555	12.01%
miR-23b-5p	TGGGTTCCTGGCATGCTGATTT	5	0.00%	2	0.00%	3	0.01%
miR-24-3p	TGGCTCAGTTCAGCAGGAACAGT	134	0.08%	100	0.08%	34	0.06%
miR-25-3p	CATTGCACTTGTCTCGGTCTGA	506	0.29%	350	0.30%	156	0.29%
miR-26a-5p	TTCAAGTAACCCAGGATAGGCT	2279	1.32%	1662	1.40%	617	1.13%
miR-26b-5p	TTCAAGTAATTCAGGATAGGTT	20	0.01%	12	0.01%	8	0.01%
miR-27a-3p	TTCACAGTGGCTAAGTTCCGC	29	0.02%	19	0.02%	10	0.02%
miR-27b-3p	TTCACAGTGGCTAAGTTCTGC	110	0.06%	84	0.07%	26	0.05%
miR-28-3p	CACTAGATTGTGAGCTCCTGGAA	127	0.07%	93	0.08%	34	0.06%
miR-28-5p	AAGGAGCTCACAGTCTATTG	12	0.01%	10	0.01%	2	0.00%
miR-29a-3p	TAGCACCATCTGAAATCGGTT	49	0.03%	31	0.03%	18	0.03%
miR-29a-5p	ACTGATTTCTTTTGGTGTTCAGA	200	0.12%	136	0.11%	64	0.12%
miR-29b-3p	TAGCACCATTTGAAATCAGTGT	45	0.03%	33	0.03%	12	0.02%
miR-29c-3p	TAGCACCATTTGAAATCGGTTA	28	0.02%	20	0.02%	8	0.01%
miR-29c-5p	TGACCGATTTCTCCTGGTGTT	515	0.30%	360	0.30%	155	0.28%
miR-30a-3p	CTTTCAGTCGGATGTTTGCAG	9	0.01%	6	0.01%	3	0.01%
miR-30a-5p	TGTAAACATCCTCGACTGGAAGC	221	0.13%	170	0.14%	51	0.09%
miR-30b-5p	TGTAAACATCCTACACTCAGCT	8	0.00%	4	0.00%	4	0.01%
miR-30c-5p	TGTAAACATCCTACACTCTCAGC	15	0.01%	10	0.01%	5	0.01%
miR-30d-5p	TGTAAACATCCCCGACTGGAAGC	1276	0.74%	978	0.83%	298	0.55%
miR-30e-5p	TGTAAACATCCTTGACTGGAAGC	180	0.10%	136	0.11%	44	0.08%
miR-32-5p	TATTGCACATTACTAAGTTGC	12	0.01%	9	0.01%	3	0.01%
miR-34a-5p	TGGCAGTGTCTTAGCTGGTTGT	29	0.02%	20	0.02%	9	0.02%
miR-34c-3p	AATCACTAACCACACGGCCAGG	23	0.01%	18	0.02%	5	0.01%
miR-92a-3p	TATTGCACTTGTCCCGGCCTGT	1779	1.03%	1077	0.91%	702	1.29%
miR-92b-3p	TATTGCACTCGTCCCGGCCTCC	195	0.11%	122	0.10%	73	0.13%
miR-92c-3p	TATTGCACTTGTCCCGGCTGT	24	0.01%	13	0.01%	11	0.02%
miR-93-5p	CAAAGTGCTGTTCGTGCAGGTAG	51	0.03%	40	0.03%	11	0.02%
miR-99a-3p	CAAGCTCGCTTCTATGGGTCT	16	0.01%	12	0.01%	4	0.01%
miR-99a-5p	AACCCGTAGATCCGATCTTGT	6236	3.61%	4574	3.87%	1662	3.04%
miR-99b-3p	CAAGCTCGTGTCTGTGGGTCCGA	24	0.01%	17	0.01%	7	0.01%
miR-99b-5p	CACCCGTAGAACCGACCTTGCG	1056	0.61%	692	0.58%	364	0.67%
miR-100-5p	AACCCGTAGATCCGAACTTGTG	1335	0.77%	1008	0.85%	327	0.60%
miR-101-3p	TACAGTACTGTGATAACTGAAT	15	0.01%	5	0.00%	10	0.02%
miR-106b-3p	CCGCACTGTGGGTACTTGCTGC	4	0.00%	2	0.00%	2	0.00%
miR-122-5p	TGGAGTGTGACAATGGTGTTT	195	0.11%	115	0.10%	80	0.15%

miR-125a-3p	ACAGGTGAGGTTCTTGGGAGCC	6	0.00%	2	0.00%	4	0.01%
miR-125a-5p	TCCCTGAGACCCTTTAACCTGT	3258	1.88%	2247	1.90%	1011	1.85%
miR-125b-3p	ACAAGTCAGGCTCTTGGGACCT	39	0.02%	27	0.02%	12	0.02%
miR-125b-5p	TCCCTGAGACCCTAACTTGTGA	30798	17.81%	22130	18.71%	8668	15.88%
miR-126-3p	TCGTACCGTGAGTAATAATGCG	5178	2.99%	3836	3.24%	1342	2.46%
miR-126-5p	CATTATTACTTTTGGTACGCG	4723	2.73%	2830	2.39%	1893	3.47%
miR-127-3p	GTTCGGATCCGTCTGAGCTTGGC	19	0.01%	18	0.02%	1	0.00%
miR-128-3p	TCACAGTGAACCGGTCTCTTT	33	0.02%	17	0.01%	16	0.03%
miR-129a-3p	AAGCCCTTACCCCAAAAAGCAT	592	0.34%	503	0.43%	89	0.16%
miR-129b-5p	CTTTTTGCGGTCTGGGCTTGC	43	0.02%	34	0.03%	9	0.02%
miR-130a-3p	CAGTGCAATGTTAAAAGGGCAT	13	0.01%	9	0.01%	4	0.01%
miR-130b-3p	CAGTGCAATGATGAAAGGGCAT	8	0.00%	5	0.00%	3	0.01%
miR-132-3p	TAACAGTCTACAGCCATGGTCGT	7	0.00%	6	0.01%	1	0.00%
miR-133a-3p	TTGGTCCCCTTCAACCAGCTGT	62	0.04%	53	0.04%	9	0.02%
miR-135-5p	TATGGCTTTTTATTCCTATGTGT	14	0.01%	12	0.01%	2	0.00%
miR-135b-5p	TATGGCTTTTCATTCCTATGTGT	6	0.00%	5	0.00%	1	0.00%
miR-139-5p	TCTACAGTGCACGTGTCTCCAGT	420	0.24%	287	0.24%	133	0.24%
miR-140-3p	TACCACAGGGTAGAACCACGGAC	58	0.03%	49	0.04%	9	0.02%
miR-140-5p	GTCAGTGGTTTTACCCTATGGTAGT	293	0.17%	233	0.20%	60	0.11%
miR-141-3p	TAACACTGTCTGGTAAAGATGG	20	0.01%	16	0.01%	4	0.01%
miR-143-3p	TGAGATGAAGCACTGTAGCTC	364	0.21%	272	0.23%	92	0.17%
miR-145-3p	ATTCCTGGAAATACTGTTCTT	20	0.01%	14	0.01%	6	0.01%
miR-145-5p	GTCCAGTTTTCCCAGGAATCCCT	5012	2.90%	4136	3.50%	876	1.60%
miR-146a-5p	TGAGAACTGAATTCCATGGGT	30	0.02%	17	0.01%	13	0.02%
miR-148a-3p	TCAGTGCACTACAGAACTTTGT	4	0.00%	1	0.00%	3	0.01%
miR-148a-5p	AAAGTTCTGAGACACTCCGACT	4	0.00%	2	0.00%	2	0.00%
miR-150-5p	GTTCTCCCAACCCTTGTACCAGT	387	0.22%	206	0.17%	181	0.33%
miR-151-3p	CTAGACTGAAGCTCCTTGAGGA	212	0.12%	139	0.12%	73	0.13%
miR-151-5p	TCGAGGAGCTCACAGTCTAGT	46	0.03%	39	0.03%	7	0.01%
miR-152-3p	TCAGTGCATGACAGAACTTGG	22	0.01%	17	0.01%	5	0.01%
miR-154b-5p	AGAGGTCTTCCATGGTGCATTCG	8	0.00%	8	0.01%	0	0.00%
miR-181a-2-3p	ACCACCGACCGTTGACTGTACC	4	0.00%	4	0.00%	0	0.00%
miR-181a-5p	AACATTCAACGCTGTCGGTGAGT	48	0.03%	39	0.03%	9	0.02%
miR-181c-3p	ACCATCGACCGTTGAGTGGACC	9	0.01%	7	0.01%	2	0.00%
miR-181c-5p	AACATTCAACCTGTCGGTGAGT	96	0.06%	72	0.06%	24	0.04%
miR-183-5p	TATGGCACTGGTAGAATTCACT	177	0.10%	123	0.10%	54	0.10%
miR-185-5p	TGGAGAGAAAGGCAGTTCCTGA	8	0.00%	5	0.00%	3	0.01%
miR-186-5p	CAAAGAATTCTCCTTTTGGGCT	96	0.06%	75	0.06%	21	0.04%
miR-187-3p	TCGTGTCTTGTGTTGCAGCCGG	5	0.00%	3	0.00%	2	0.00%
miR-191-5p	CAACGGAATCCCAAAAGCAGCT	301	0.17%	230	0.19%	71	0.13%
miR-192-5p	CTGACCTATGAATTGACAGCCAT	77	0.04%	35	0.03%	42	0.08%
miR-193a-3p	AACTGGCCTACAAAGTCCCAGT	17	0.01%	10	0.01%	7	0.01%
miR-193a-5p	TGGGTCTTTGCGGGCGAGATGA	670	0.39%	480	0.41%	190	0.35%
miR-193b-3p	AACTGGCCCACAAAGTCCCGCT	121	0.07%	90	0.08%	31	0.06%
miR-193b-5p	CGGGGTTTTGAGGGCGAGATGA	5	0.00%	5	0.00%	0	0.00%
miR-195-5p	TAGCAGCACAGAAATATTGGC	54	0.03%	42	0.04%	12	0.02%
miR-196b-5p	TAGGTAGTTTCCTGTTGTTGG	58	0.03%	42	0.04%	16	0.03%
miR-199a-3p	CAGTAGTCTGCACATTGGTTAA	34	0.02%	27	0.02%	7	0.01%
miR-199a-5p	CCCAGTGTTCAGACTACCTGTT	82	0.05%	54	0.05%	28	0.05%
miR-199b-3p	ACAGTAGTCTGCACATTGGTT	220	0.13%	160	0.14%	60	0.11%
miR-199b-5p	CCCAGTGTTTAGACTATCTGTTC	5	0.00%	4	0.00%	1	0.00%
miR-200a-3p	TAACACTGTCTGGTAACGATGTT	460	0.27%	288	0.24%	172	0.32%
miR-200a-5p	CATCTTACCGGACAGTGCTGGA	39	0.02%	25	0.02%	14	0.03%
miR-200b-3p	TAATACTGCCTGGTAATGATGA	18150	10.50%	11247	9.51%	6903	12.64%
miR-200c-3p	TAATACTGCCGGGTAATGATGGA	18733	10.83%	11559	9.77%	7174	13.14%
miR-204-5p	TTCCCTTTGTCATCCTATGCCT	45	0.03%	36	0.03%	9	0.02%
miR-210-3p	CTGTGCGTGTGACAGCGGCTGA	4	0.00%	4	0.00%	0	0.00%
miR-214-3p	TACAGCAGGCACAGACAGGCAGT	20	0.01%	20	0.02%	0	0.00%
miR-214-5p	TGCCTGTCTACACTTGCTGTGC	149	0.09%	108	0.09%	41	0.08%
miR-218b-5p	TTGTGCTTGATCTAACCATGT	365	0.21%	254	0.21%	111	0.20%
miR-221-3p	AGCTACATTGTCTGCTGGGTTTC	158	0.09%	114	0.10%	44	0.08%
miR-222-3p	AGCTACATCTGGCTACTGGGTCT	26	0.02%	20	0.02%	6	0.01%
miR-299-5p	TGGTTTACCGTCCCACATACAT	12	0.01%	9	0.01%	3	0.01%
miR-320-3p	AAAAGCTGGGTTGAGAGGGCGA	168	0.10%	121	0.10%	47	0.09%
miR-324-3p	ACTGCCCCAGGTGCTGCTGGT	118	0.07%	84	0.07%	34	0.06%
miR-324-5p	CGCATCCCCTAGGGCATTGGTGT	69	0.04%	50	0.04%	19	0.03%
miR-328-3p	CTGGCCCTCTCTGCCCTTCCGT	26	0.02%	17	0.01%	9	0.02%

miR-329-3p	AACACACCTGGTTAACCTCTTT	17	0.01%	8	0.01%	9	0.02%
miR-331-5p	TCTAGGTATGGTCCCAGGGATC	17	0.01%	10	0.01%	7	0.01%
miR-338-5p	AACAATATCCTGGTGCTGAGT	5	0.00%	1	0.00%	4	0.01%
miR-339-3p	AGCTCCTCGAGGCCAGAGCC	10	0.01%	9	0.01%	1	0.00%
miR-339-5p	TCCCTGTCCTCCAGGAGCTCA	91	0.05%	59	0.05%	32	0.06%
miR-340-3p	TCCGTCTCAGTTACTTTATAGCC	6	0.00%	6	0.01%	0	0.00%
miR-345-3p	CCTGAACTAGGGGTCTGGAGT	6	0.00%	3	0.00%	3	0.01%
miR-345-5p	GCTGACTCCTAGTCCAGTGCT	8	0.00%	6	0.01%	2	0.00%
miR-361-5p	TTATCAGAATCTCCAGGGGTAC	23	0.01%	18	0.02%	5	0.01%
miR-362-3p	AACACACCTATTCAAGGATTCA	5	0.00%	3	0.00%	2	0.00%
miR-362-5p	AATCCTTGGAACCTAGGTGTGAGT	211	0.12%	148	0.13%	63	0.12%
miR-365-3p	TAATGCCCCTAAAAATCCTTAT	632	0.37%	425	0.36%	207	0.38%
miR-374a-5p	TTATAATACAACCTGATAAGTGT	425	0.25%	260	0.22%	165	0.30%
miR-374b-5p	ATATAATACAACCTGCTAAGTGT	84	0.05%	50	0.04%	34	0.06%
miR-375-3p	TTTGTTCGTTCGGCTCGCGTGA	98	0.06%	73	0.06%	25	0.05%
miR-377-5p	AGAGGTTGCCCTTGGTGAATTC	9	0.01%	8	0.01%	1	0.00%
miR-378-3p	ACTGGACTTGGAGTCAGAAGGC	2847	1.65%	2254	1.91%	593	1.09%
miR-378-5p	CTCCTGACTCCAGGTCCTGTGT	48	0.03%	38	0.03%	10	0.02%
miR-409-3p	GAATGTTGCTCGGTGAACCCCT	42	0.02%	30	0.03%	12	0.02%
miR-409-5p	AGGTTACCCGAGCAACTTTGCA	4	0.00%	2	0.00%	2	0.00%
miR-423-5p	TGAGGGGCAGAGAGCGAGAC	189	0.11%	133	0.11%	56	0.10%
miR-425-3p	CATCGGGAATGTCGTGTCCGCC	25	0.01%	19	0.02%	6	0.01%
miR-425-5p	AATGACACGATCACTCCCGTTGA	5	0.00%	4	0.00%	1	0.00%
miR-429-3p	TAATACTGTCTGGTAATGCCGT	148	0.09%	86	0.07%	62	0.11%
miR-450b-5p	TTTTGCAATATGTTCCTGAAT	9	0.01%	7	0.01%	2	0.00%
miR-451-5p	AAACCGTTACCATTACTGAGTTT	342	0.20%	238	0.20%	104	0.19%
miR-455-3p	GCAGTCCATGGGCATATACACT	5	0.00%	4	0.00%	1	0.00%
miR-455-5p	TATGTGCCTTTGGACTACATCG	191	0.11%	126	0.11%	65	0.12%
miR-483-3p	TCACTCCTCTCCTCCCGTCTTCC	5	0.00%	2	0.00%	3	0.01%
miR-484-5p	TCAGGCTCAGTCCCCTCCCGAT	12	0.01%	7	0.01%	5	0.01%
miR-485-5p	AGAGGCTGGCCGTGATGAATTC	9	0.01%	5	0.00%	4	0.01%
miR-486-3p	TCCTGTACTGAGCTGCCCCGAG	70	0.04%	44	0.04%	26	0.05%
miR-497-5p	CAGCAGCACACTGTGGTTTGT	12	0.01%	7	0.01%	5	0.01%
miR-500-3p	ATGCACCTGGGCAAGGATTCTGA	31	0.02%	21	0.02%	10	0.02%
miR-500-5p	TAATCCTTGCTACCTGGGTGAGA	273	0.16%	152	0.13%	121	0.22%
miR-503-5p	TAGCAGCGGGAACAGTACTGCA	4	0.00%	2	0.00%	2	0.00%
miR-504-5p	AGACCCTGGTCTGCACTCTATC	33	0.02%	27	0.02%	6	0.01%
miR-505-3p	GTCAACACTTGCTGGTTTCCTCT	8	0.00%	4	0.00%	4	0.01%
miR-532-3p	GTCCTCCCACACCCAAGGCTTGCA	151	0.09%	82	0.07%	69	0.13%
miR-532-5p	CATGCCTTGAGTGTAGGACCGT	172	0.10%	122	0.10%	50	0.09%
miR-539-5p	GGAGAAATTATCCTTGGTGTGT	53	0.03%	29	0.02%	24	0.04%
miR-551a-3p	GTGCGACCCACTCTTGGTTTCCA	18	0.01%	12	0.01%	6	0.01%
miR-551b-3p	GCGACCCATACTTGGTTTCAGT	6	0.00%	6	0.01%	0	0.00%
miR-574-3p	CACGCTCATGCACACACCCACA	37	0.02%	27	0.02%	10	0.02%
miR-652-3p	AATGGCGCCACTAGGGTTGTGT	43	0.02%	34	0.03%	9	0.02%
miR-664-5p	CAGGCTAGGAGAAGTGATTGGA	54	0.03%	37	0.03%	17	0.03%
miR-670-5p	ATCCCTGAGTGTATGTGGTGA	25	0.01%	23	0.02%	2	0.00%
miR-676-3p	CCGTCCTAAGGTTGTTGAGTTT	208	0.12%	151	0.13%	57	0.10%
miR-758-3p	TTTGTGACCTGGTCCACTAAC	19	0.01%	9	0.01%	10	0.02%
miR-769-5p	TGAGACCTCTGGGTTCTGAGCT	24	0.01%	15	0.01%	9	0.02%
miR-874-3p	CTGCCCTGGCCCGAGGGACCGA	35	0.02%	23	0.02%	12	0.02%
miR-1280-5p	ATCCCACCGCTGCCACCA	4	0.00%	3	0.00%	1	0.00%
miR-1468-5p	CTCCGTTTGCCTGTTTTGCTGA	41	0.02%	28	0.02%	13	0.02%
miR-1983-3p	CTCACCTGGAGCATGTTTTCT	11	0.01%	9	0.01%	2	0.00%
miR-2483-3p	AAACATCTGGTTGGTTGAGAGA	11	0.01%	9	0.01%	2	0.00%
miR-2779-5p	TCCGGCTCGAAGGACCA	33	0.02%	27	0.02%	6	0.01%
miR-2898-3p	TCCCCGGCATCTCCACCA	44	0.03%	33	0.03%	11	0.02%
miR-4286-5p	ACCCCACTCCTGGTACCA	84	0.05%	63	0.05%	21	0.04%
miR-4454-5p	TCGAATCCGAGTCACGGCACCA	24	0.01%	19	0.02%	5	0.01%
miR-4497-5p	CGGCTCCGGGACGGCTGGGA	7	0.00%	4	0.00%	3	0.01%
miR-5100-3p	TCGAATCCCAGCGGTGCCTCCA	29	0.02%	22	0.02%	7	0.01%

[a]Sequence with the major sequence copy number; [b]CN = the sequence copy number.

Table 5. RT-qPCR expression level of biological candidate miRNAs in pregnant uterus of 35 IbxMe F_2 sows regarding prolificacy levels. miRNAs expression was measured in 35 IbxMe F_2 sows that were classified into two groups according the NE on day 30-32 of gestation: high (n = 16; NE ≥ 13 and low (n=19; NE ≤ 11).

miRNA	Mean High Prolificacy Sows	Mean Low Prolificacy Sows	Fold Change High/Low[a]	*P-Value*[b]
miR-139-5p	1.79	2.22	-1.24	0.02
miR-150-5p	2.36	2.99	-1.27	0.05
miR-27a-3p	1.97	2.30	-1.17	0.08
miR-20-5p	3.97	5.03	-1.27	0.09
miR-21-5p	5.88	4.95	1.19	0.18
miR-101-3p	9.48	7.76	1.22	0.41
miR-122-5p	22.14	19.20	1.15	0.67
miR-199b-3p	2.69	2.67	1.01	0.94

[a]Positive and negative signs indicate whether the level of gene expression is higher in the first or in the second group, respectively.

[b]Nominal *P-Value*. Significance threshold was set at 0.05.

Table 6. Gene expression profile of putative targets for prolificacy-related miRNAs and their expression correlation. Expression correlations were measured in uterus of 14 IbxMe F_2 pregnant sows regarding prolificacy levels. For mRNA targets and prolificacy-related miRNAs, microarray and RT-qPCR data was used, respectively.

Affymetrix Probe ID	Gene Name	SSC[a]	Log Fold Change High/Low	Adjusted *P-Value*	miR 20a-5p[b]	miR 27a-3p[b]	miR 150-5p[b]	miR 139-5p[b]
Ssc.7146.1.A1_at	*ABCA1*	1	2.102	0.026	-0.350	**-0.540**		-0.400
Ssc.8122.1.A1_at	*ACACA*	12	-2.660	0.003		0.410		
Ssc.941.1.S1_at	*ACTA2*	14	-2.408	0.002		**0.710**		
Ssc.8904.1.A1_at	*ADAMTSL3*	7	-2.953	0.003		**0.660**		
Ssc.9176.1.A1_at	*AEBP2*	5	-2.063	0.008	0.170			0.440
Ssc.29811.1.A1_at	*AKAP2*	1	-2.014	0.010			-0.190	
Ssc.15403.1.S1_at	*AMOTL1*	9	-2.768	0.004		**0.720**	-0.200	
Ssc.4135.2.A1_at	*AP1S2*	x	2.442	0.037				**-0.560**
Ssc.4033.2.S1_at	*ARSE*		2.825	0.001	-0.290			
Ssc.29004.1.S1_at	*BMP2*	17	2.657	0.002	-0.220			
Ssc.18080.1.A1_at	*C8orf46*	4	-2.113	0.036			-0.430	
Ssc.17313.1.A1_at	*CALD1*	18	-2.189	0.006	0.370	**0.700**		
Ssc.29135.1.S1_at	*CAMK2A*		-2.511	0.004		**0.640**		
Ssc.2491.1.S1_at	*CAMK2G*	14	-2.074	0.002			-0.140	
Ssc.8079.1.S1_at	*CAMK2N1*	6	-2.285	0.002	0.440			
Ssc.5621.1.S1_at	*CAP2*	7	-3.841	0.001		**0.690**		
Ssc.19533.1.S1_at	*CCDC93*	15	-3.426	0.001	0.190			
Ssc.7106.1.S1_at	*CDO1*	2	-2.214	0.047				**0.620**
Ssc.7645.1.A1_at	*CFL2*	7	-3.640	0.004	0.470	**0.710**		
Ssc.12523.1.A1_at	*CLIC4*		-2.089	0.016	0.420			0.250
Ssc.9013.1.S1_at	*CNN1*	2	-4.621	0.001	0.340			
Ssc.4779.1.A1_at	*CPM*		2.096	0.026				-0.500
Ssc.4984.1.S1_at	*CXCL14*	2	2.930	0.031	-0.110		-0.160	
Ssc.22436.1.S1_at	*CYP26B1*	3	2.342	0.007	-0.260			
Ssc.11310.2.A1_at	*CHRDL1*	x	-3.084	0.005		**0.790**		
Ssc.29980.1.S1_at	*DCLK1*	11	-2.046	0.005				**0.554**
Ssc.27562.1.S1_at	*DDX5*		-3.671	0.003	0.350			
Ssc.25324.1.S1_at	*DMD*	X	-3.520	0.002				0.460
Ssc.19368.1.S1_at	*DPP6*	18	-3.147	0.017				0.420
Ssc.31073.1.S1_at	*DPYD*	4	-2.632	0.009	0.280			
Ssc.1623.1.S1_at	*DSP*	7	2.079	0.028	-0.230	**-0.550**		
Ssc.29763.1.A1_at	*ENDOU*		-2.538	0.001		**0.590**		**0.588**
Ssc.83.1.S1_at	*ENPEP*	8	-2.093	0.048		**0.560**		
Ssc.19907.1.S1_at	*F3*	4	2.023	0.034	-0.270	**-0.660**		
Ssc.21691.3.S1_at	*FADS6*	12	2.887	0.004			-0.140	
Ssc.998.1.A1_at	*FAM129A*	9	-2.014	0.034	0.170			
Ssc.27703.1.S1_at	*FBN2*	2	-2.543	0.001				0.530
Ssc.639.1.A1_at	*GEM*	4	-3.435	0.003		**0.700**		
Ssc.29187.1.A1_at	*GJA1*	1	-2.070	0.039	0.180			
Ssc.18554.1.S1_x_at	*HLA-B*	7	2.468	0.003			0.270	
Ssc.12578.1.A1_at	*IGF1*	5	-2.991	0.003		**0.640**		
Ssc.7864.1.A1_at	*IL1RAP*	13	2.204	0.017	-0.130	-0.440		
Ssc.23179.1.A1_at	*INPP5A*	14	-2.194	0.001			-0.140	
Ssc.16218.1.S1_at	*ITGA2*	16	2.609	0.011		**-0.630**		**-0.650**
Ssc.16663.1.S1_at	*ITGA5*	5	-2.373	0.008		**0.640**		
Ssc.27409.1.S1_at	*JPH2*	17	-2.082	0.003			-0.190	

Ssc.2444.2.A1_a_at	*LAMA3*	6	2.239	0.010	-0.100			
Ssc.21806.1.A1_at	*LHFP*	2	-2.191	0.003	0.420	**0.720**		
Ssc.4274.1.S1_at	*LRP1B*	15	-2.248	0.014	0.330			
Ssc.29636.1.A1_at	*MAMDC2*	1	-2.625	0.018			-0.170	
Ssc.9434.1.A1_at	*MARK1*	10	2.982	0.004		-0.490		**-0.590**
Ssc.19455.1.S1_at	*MGLL*	13	-2.749	0.002	0.380			
Ssc.17375.1.S1_at	*MRAP2*	1	-2.561	0.002			-0.030	
Ssc.12367.1.A1_at	*MUC4*	13	3.337	0.014			-0.040	
Ssc.19282.1.S1_at	*MXRA7*		-2.188	0.001				0.48
Ssc.16164.1.S1_at	*MYLK*	13	-2.311	0.003				0.527
Ssc.12431.1.A1_at	*MYO5B*	1	3.602	0.014	-0.150			
Ssc.6156.1.A1_at	*NBL1*	6	-2.111	0.003	0.410			
Ssc.19592.3.S1_a_at	*NCS1*	1	-2.560	0.001		**0.630**	-0.130	
Ssc.24221.2.A1_at	*NDUFS4*	16	-3.643	0.005		**0.780**		
Ssc.17615.1.S1_at	*NME7*	7	2.290	0.013				**-0.540**
Ssc.8125.1.A1_at	*NPTXR*	5	-2.214	0.010		**0.620**		
Ssc.19579.2.S1_at	*NR2F2*	7	-2.015	0.005		**0.550**	-0.280	
Ssc.7463.1.A1_at	*PALLD*	14	-2.380	0.001	0.340			
Ssc.10303.1.A1_at	*PDGFD*	9	-2.116	0.013	0.510			
Ssc.12977.1.A1_at	*PDLIM4*	2	-2.079	0.003	0.260			
Ssc.526.1.S1_at	*PKIA*	4	-3.059	0.004	0.340	**0.650**		
Ssc.10743.1.A1_at	*PKP2*	5	2.173	0.001		**-0.540**		
Ssc.27627.1.S1_at	*PPP1R12B*	10	-2.293	0.006	0.130			
Ssc.19839.1.S1_at	*PSAT1*	1	2.009	0.025				**-0.620**
Ssc.9991.1.S1_at	*PTHLH*	5	3.570	0.017		**-0.540**		
Ssc.24481.1.S1_at	*RASL12*		-2.847	0.002	0.430			
Ssc.11787.2.A1_at	*RASSF2*	17	-2.622	0.002	0.370			
Ssc.9607.1.A1_at	*RCAN1*	13	-2.486	0.001			-0.080	
Ssc.24238.1.S1_s_at	*RECQL5*	12	2.062	0.027			0.090	
Ssc.30886.1.S1_at	*RNGTT*	1	-2.059	0.021		**0.690**		0.440
Ssc.18217.1.S1_at	*RSPO3*	1	-2.207	0.026		**0.800**		
Ssc.6765.1.S1_at	*SCIN*	9	2.332	0.017			-0.060	
Ssc.6666.1.A1_at	*SCHIP1*	13	-3.518	0.002				0.520
Ssc.2283.1.S1_at	*SFRP1*	17	-2.634	0.001		**0.660**		
Ssc.14331.1.A1_at	*SIK3*	9	-2.672	0.012	0.270			
Ssc.4483.1.A1_at	*SLC19A3*		3.200	0.028			0.200	
Ssc.4705.2.A1_at	*SMOC2*	1	-2.438	0.008	0.420	**0.730**		
Ssc.10920.1.S1_a_at	*SYNM*	1	-4.913	0.002	0.320			
Ssc.26929.1.S1_at	*SYNPO2*	8	-3.309	0.007	0.460			0.450
Ssc.16704.1.S1_at	*TACC1*	15	-2.430	0.002	0.350			
Ssc.2675.1.S1_at	*TGFB1I1*		-2.568	0.003	0.270			
Ssc.992.1.S1_at	*THBS2*	1	-2.180	0.004	0.290			
Ssc.2971.1.S1_at	*TNS1*	15	-2.357	0.001	0.330		-0.190	0.440
Ssc.17874.1.A1_at	*TOX*	4	-2.292	0.008		**0.690**		**0.690**
Ssc.17490.1.S1_at	*TRDN*	1	2.080	0.050	-0.320	**-0.600**		
Ssc.19067.1.A1_at	*WDR17*	15	-2.722	0.004			-0.210	
Ssc.9299.1.S1_at	*ZEB1*	10	-2.016	0.004			-0.120	
Ssc.25186.1.A1_at	*ZFPM2*	4	-2.359	0.013	0.460			

[a] Location of genes in Sus Scrofa Chromosome (SSC) (S.Scrofa 10.2).

[b] Pearson correlation. Significance for bold values was set at 0.05.

Table 7. Differentially expressed genes in uterus regarding prolificacy levels and putative targets for prolificacy-related miRNAs which are proposed to be candidate genes for significant QTL on litter size (NBA and TNB) (Noguera *et al.*, 2009,Balcells *et al.*, 2011).

SSC[a]	QTL (cM)	Trait[1]	Candidate Genes	Prolificacy-related miRNAs	Reference
1	153 (146-153)	NBA	*AKAP2*	miR-150-5p	(Noguera *et al.*, 2009)
	139 (132-147)	TNB	*NCS1*	miR-150-5p, miR-27a-3p	
1	76 (69-84)	TNB	*RNGTT*	miR-139-5p, miR-27a-3p	
	79 (72-84)	NBA	*MYO5B*	miR-20a-5p	
4	23(21-26)	TNB	*GEM*	miR-27a-3p	
			PKIA	miR-20a-5p, miR-27a-3p	
5	66(59-73)	NBA	*ITGA5*	miR-27a-3p	
			PKP2	miR-27a-3p	
			PTHLH	miR-27a-3p	
			AEBP2	miR-150-5p, miR-139-5p	
6	4(1-10)	NBA	*NBL1*	miR-20a-5p	
			CAMK2N1	miR-20a-5p	
7	24(12-33)	TNB	*DSP*	miR-20a-5p, miR-27a-3p	
	28(12-37)		*CAP2*	miR-27a-3p	
	70(46-77)		*HLA-B*	miR-150-5p	
7	89(84-96)	TNB	*ADAMTSL3*	miR-27a-3p	
			CFL2	miR-20a-5p, miR-27a-3p	
	107(100-116)	TNB/NBA	*NR2F2*	miR-150-5p, miR-27a-3p	
			NME7	miR-139-5p	
8	83(76-89)	TNB	*SYNPO2*	miR-139-5p, miR-20a-5p	
	92(88-94)	NBA			
12	5(1-10)	TNB	*RECQL5*	miR-150-5p	
	11(8-19)	TNB	*FADS6*	miR-150-5p	
	11(9-18)	NBA			
14	29(25-36)	NBA	*PALLD*	miR-20a-5p	
15	3(1-8)	NBA	*LRP1B*	miR-20a-5p	
18	11(2-20)	NBA	*CALD1*	miR-20a-5p, miR-27a-3p	
13	44(22-77)	NBA	*MGLL*	miR-20a-5p	(Balcells *et al.*, 2011)
			SCHIP1	miR-139-5p	
	51(33-74)	TNB	*IL1RAP*	miR-20a-5p, miR-27a-3p	
			MUC4	miR-150-5p	
			MYLK	miR-139-5p	
17	22(11-42)	NBA	*BMP2*	miR-20a-5p	(Noguera *et al.*, 2009)
	22(12-62)	TNB	*JPH2*	miR-150-5p	

[a] Porcine chromosome (SSC)

[b] NBA = Number of piglets born alive; TNB = Total number of piglets born.

FIGURES

Figure 1. Plot of the log-fold change against the log-concentration for each miRNA (188 mature miRNAs) expressed in pregnant uterus. miRNA expression profile was evaluated in 14 IbxMe F_2 regarding prolificacy levels according the NE on day 30-32 of gestation (high (n = 5; NE ≥ 14) and low (n = 9; NE ≤ 11)).

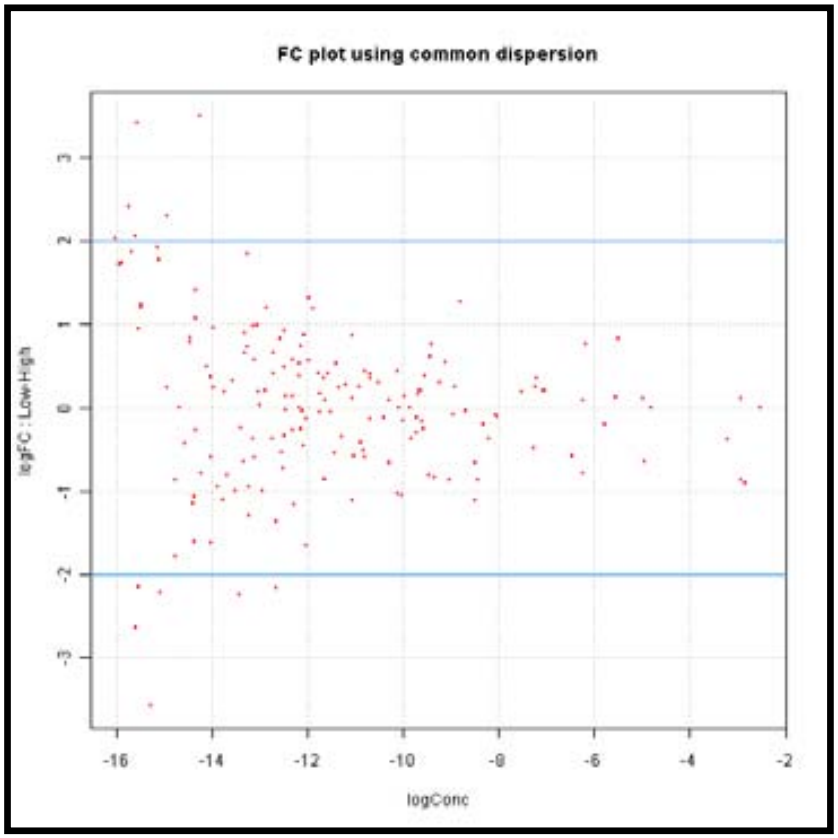
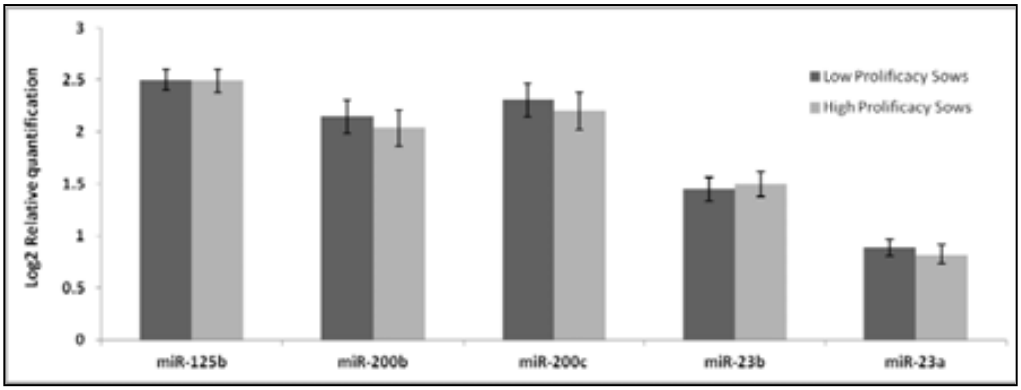


Figure 2. Relative quantification of the most abundant miRNAs (miR-125b-5p, miR-23b-3p, miR-200c-3p, miR-200b-3p and miR-23a-3p) in uterus of pregnant sows. miRNAs expression was measured in 35 IbxMe F_2 sows that were classified into two groups according the NE on day 30-32 of gestation: high (n = 16; NE ≥ 13 and low (n=19; NE ≤ 11).



SUPPLEMENTARY TABLE

Supplementary Table 1. Microarray expression profile of differential expressed genes (Log Fold Change >2) in uterus of IbxMe F_2 sows regarding prolificacy levels.

ID	Gene Name	Gene Product	Log Fold Change High/Low[a]	Adjusted *P-Value*[b]
Ssc.3835.1.S1_at	*PDLIM3*	PDZ and LIM domain 3	-5.98	0.0007
Ssc.10064.2.A1_at	*KCNMB1*	potassium large conductance calcium-activated channel, subfamily M, beta member 1	-5.20	0.0007
Ssc.10316.1.S1_at	*ACTG2*	actin, gamma 2, smooth muscle, enteric	-5.03	0.0013
Ssc.10920.1.S1_a_at	*SYNM*	synemin, intermediate filament protein	-4.91	0.0022
Ssc.24.1.S1_at	*DES*	desmin	-4.76	0.0007
Ssc.19476.2.A1_s_at	*PYGM*	phosphorylase, glycogen, muscle	-4.68	0.0000
Ssc.9013.1.S1_at	*CNN1*	calponin 1, basic, smooth muscle	-4.62	0.0007
Ssc.18344.1.S1_at	*PYGM*	phosphorylase, glycogen, muscle	-4.54	0.0000
Ssc.27427.1.S1_a_at	*SORCS1*	sortilin-related VPS10 domain containing receptor 1	-4.38	0.0007
Ssc.23294.1.S1_at	*SYNM*	synemin, intermediate filament protein	-4.24	0.0014
Ssc.22072.2.S1_at	*TSPAN2*	tetraspanin 2	-4.14	0.0009
Ssc.11281.1.A1_at	*PENK*	proenkephalin	-4.13	0.0007
Ssc.415.1.S1_at	*CKM*	creatine kinase, muscle	-4.08	0.0007
Ssc.5227.1.S2_at	-	No annotation	-4.06	0.0054
Ssc.27727.2.S1_a_at	*PGM5*	phosphoglucomutase 5	-4.06	0.0007
Ssc.5227.1.S1_at	*PLN*	phospholamban	-4.05	0.0024
Ssc.3232.1.S1_at	*SFRP2*	secreted frizzled-related protein 2	-3.99	0.0006
Ssc.14368.1.A1_at	*MMRN1*	multimerin 1	-3.97	0.0017
Ssc.8683.1.S1_at	-	No annotation	-3.90	0.0040
Ssc.16209.1.S1_at	*TNC*	tenascin C	-3.85	0.0007
Ssc.29185.1.A1_at	*ITGA8*	integrin, alpha 8	-3.84	0.0116
Ssc.5621.1.S1_at	*CAP2*	CAP, adenylate cyclase-associated protein, 2 (yeast)	-3.84	0.0014
Ssc.14463.1.S1_at	*FHL1*	four and a half LIM domains 1	-3.82	0.0003
Ssc.20571.1.S1_a_at	*TPM2*	tropomyosin 2 (beta)	-3.80	0.0010
Ssc.23247.1.S1_at	*MYLK*	myosin light chain kinase	-3.73	0.0032
Ssc.30718.1.A1_at	*KCNH2*	potassium voltage-gated channel, subfamily H (eag-related), member 2	-3.71	0.0008
Ssc.20571.2.S1_a_at	*TPM2*	tropomyosin 2 (beta)	-3.70	0.0013
Ssc.27562.1.S1_at	*DDX5*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	-3.67	0.0035
Ssc.27727.1.S1_at	*PGM5*	phosphoglucomutase 5	-3.65	0.0007
Ssc.1534.1.A1_at	*SORBS1*	sorbin and SH3 domain containing 1	-3.65	0.0011
Ssc.24221.2.A1_at	*NDUFS4*	NADH dehydrogenase (ubiquinone) Fe-S protein 4, 18kDa (NADH-coenzyme Q reductase)	-3.64	0.0054
Ssc.7645.1.A1_at	*CFL2*	cofilin 2 (muscle)	-3.64	0.0036
Ssc.10391.1.A1_at	*MCAM*	melanoma cell adhesion molecule	-3.61	0.0007
Ssc.10392.1.A1_at	*MYH11*	myosin, heavy chain 11, smooth muscle	-3.60	0.0010
Ssc.28465.1.S1_at	*PGM5*	phosphoglucomutase 5	-3.60	0.0007
Ssc.7645.2.S1_at	*CFL2*	cofilin 2 (muscle)	-3.57	0.0042
Ssc.4076.1.S1_at	*RBKS*	ribokinase	-3.57	0.0013
Ssc.25324.1.S1_at	*DMD*	dystrophin	-3.52	0.0024
Ssc.6666.1.A1_at	*SCHIP1*	schwannomin interacting protein 1	-3.52	0.0021
Ssc.28696.1.S1_at	*DMPK*	dystrophia myotonica-protein kinase	-3.50	0.0006
Ssc.12842.1.S2_at	*CAV1*	caveolin 1, caveolae protein, 22kDa	-3.49	0.0020
Ssc.639.1.A1_at	*GEM*	GTP binding protein overexpressed in skeletal muscle	-3.44	0.0032
Ssc.19533.1.S1_at	*Q8TBX5*	No annotation	-3.43	0.0013

Ssc.934.1.S1_at	*NEXN*	nexilin (F actin binding protein)	-3.42	0.0048
Ssc.10391.2.S1_at	*MCAM*	melanoma cell adhesion molecule	-3.41	0.0007
Ssc.28305.1.A1_at	*TACC1*	transforming, acidic coiled-coil containing protein 1	-3.36	0.0175
Ssc.9125.1.S1_at	*MYL9*	myosin, light chain 9, regulatory	-3.33	0.0007
Ssc.25004.1.S1_at	*IGF1*	insulin-like growth factor 1 (somatomedin C)	-3.32	0.0028
Ssc.26929.1.S1_at	*SYNPO2*	synaptopodin 2	-3.31	0.0072
Ssc.1377.2.S1_at	*ITGA8*	integrin, alpha 8	-3.31	0.0060
Ssc.31130.1.A1_at	*IGF1*	insulin-like growth factor 1 (somatomedin C)	-3.30	0.0016
Ssc.8405.1.A1_at	*PCDH15*	protocadherin 15	-3.29	0.0048
Ssc.26180.1.S1_at	*OGN*	osteoglycin	-3.24	0.0089
Ssc.1169.1.A1_s_at	*MFAP4*	microfibrillar-associated protein 4	-3.22	0.0007
Ssc.5883.1.S1_at	*TTC7A*	tetratricopeptide repeat domain 7A	-3.21	0.0003
Ssc.9452.1.S1_at	*MYH11*	myosin, heavy chain 11, smooth muscle	-3.16	0.0036
Ssc.19368.1.S1_at	*DPP6*	dipeptidyl-peptidase 6	-3.15	0.0165
Ssc.29226.1.S1_at	*DCLK1*	doublecortin-like kinase 1	-3.12	0.0032
Ssc.1411.1.S1_at	*THBS4*	thrombospondin 4	-3.10	0.0032
Ssc.11310.2.A1_at	*CHRDL1*	chordin-like 1	-3.08	0.0053
Ssc.3343.1.S1_at	*NP_057709*	No annotation	-3.08	0.0033
Ssc.526.1.S1_at	*PKIA*	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	-3.06	0.0039
Ssc.1716.1.S1_at	*ASPN*	asporin	-3.03	0.0177
Ssc.11913.1.S1_at	*TAGLN*	transgelin	-3.02	0.0008
Ssc.22072.1.A1_at	*TSPAN2*	tetraspanin 2	-3.02	0.0029
Ssc.6219.1.A1_at	*GLDN*	gliomedin	-3.02	0.0038
Ssc.22100.1.S1_at	*SH3BGR*	SH3 domain binding glutamic acid-rich protein	-3.01	0.0009
Ssc.12578.1.A1_at	*IGF1*	insulin-like growth factor 1 (somatomedin C)	-2.99	0.0032
Ssc.373.1.S1_at	*MGP*	matrix Gla protein	-2.98	0.0044
Ssc.7552.1.A1_at	*TES*	testis derived transcript (3 LIM domains)	-2.98	0.0011
Ssc.22761.1.S1_at	*CFL2*	cofilin 2 (muscle)	-2.97	0.0013
Ssc.8904.1.A1_at	*ADAMTSL3*	ADAMTS-like 3	-2.95	0.0030
Ssc.9236.1.A1_at	*TPM2*	tropomyosin 2 (beta)	-2.90	0.0039
Ssc.15710.1.A1_at	*ECHDC1*	enoyl Coenzyme A hydratase domain containing 1	-2.90	0.0011
Ssc.4661.1.S1_at	*C1orf41*	No Annotats GO David	-2.88	0.0054
Ssc.3621.1.S1_at	*DPYSL3*	dihydropyrimidinase-like 3	-2.88	0.0002
Ssc.3630.1.S1_at	*LIMS2*	LIM and senescent cell antigen-like domains 2	-2.87	0.0013
Ssc.11310.1.A1_at	*ANK3*	ankyrin 3, node of Ranvier (ankyrin G)	-2.85	0.0047
Ssc.24481.1.S1_at	*RASL12*	RAS-like, family 12	-2.85	0.0016
Ssc.21758.1.S1_at	*FLNA*	filamin A, alpha (actin binding protein 280)	-2.85	0.0007
Ssc.25882.1.S1_at	*SSPN*	sarcospan (Kras oncogene-associated gene)	-2.83	0.0036
Ssc.266.1.S1_at	-	No annotation	-2.83	0.0098
Ssc.26452.1.S1_at	*PDZD4*	PDZ domain containing 4	-2.83	0.0106
Ssc.24221.1.S1_at	*BIRC6*	baculoviral IAP repeat-containing 6	-2.83	0.0034
Ssc.1592.1.S1_at	*DPYSL3*	dihydropyrimidinase-like 3	-2.83	0.0008
Ssc.1170.1.A1_at	*MSRB3*	methionine sulfoxide reductase B3	-2.82	0.0056
Ssc.11945.1.A1_at	*SGCD*	sarcoglycan, delta (35kDa dystrophin-associated glycoprotein)	-2.81	0.0013
Ssc.2207.2.S1_a_at	*SMTN*	smoothelin	-2.81	0.0011
Ssc.1377.3.S1_at	*ITGA8*	integrin, alpha 8	-2.79	0.0186
Ssc.11770.1.S1_at	*PDLIM7*	PDZ and LIM domain 7 (enigma)	-2.78	0.0010
Ssc.15403.1.S1_at	*AMOTL1*	angiomotin like 1	-2.77	0.0038
Ssc.19476.1.S1_at	*PYGM*	phosphorylase, glycogen, muscle	-2.76	0.0013
Ssc.29492.1.A1_at	*ANK3*	ankyrin 3, node of Ranvier (ankyrin G)	-2.75	0.0053
Ssc.12842.1.S1_at	*CAV1*	caveolin 1, caveolae protein, 22kDa	-2.75	0.0030
Ssc.19455.1.S1_at	*MGLL*	monoglyceride lipase	-2.75	0.0017

Ssc.19067.1.A1_at	*WDR17*	WD repeat domain 17	-2.72	0.0043
Ssc.4204.1.S1_at	*GYG1*	glycogenin 1	-2.72	0.0032
Ssc.24075.1.A1_at	*DCLK1*	doublecortin-like kinase 1	-2.72	0.0123
Ssc.16589.1.S1_at	*COL6A3*	collagen, type VI, alpha 3	-2.70	0.0007
Ssc.30224.1.A1_at	*ARHGAP15*	Rho GTPase activating protein 15	-2.69	0.0035
Ssc.4118.2.S1_at	*C20orf123*	chromosome 20 open reading frame 123	-2.69	0.0041
Ssc.18613.1.S1_at	*CCL21*	chemokine (C-C motif) ligand 21	-2.68	0.0009
Ssc.14331.1.A1_at	*Q9Y2K2*	No annotation	-2.67	0.0120
Ssc.8122.1.A1_at	*ACACA*	acetyl-Coenzyme A carboxylase alpha	-2.66	0.0032
Ssc.21769.1.S1_at	*RERG*	RAS-like, estrogen-regulated, growth inhibitor	-2.64	0.0061
Ssc.2283.1.S1_at	*SFRP1*	secreted frizzled-related protein 1	-2.63	0.0013
Ssc.31073.1.S1_at	*DPYD*	dihydropyrimidine dehydrogenase	-2.63	0.0090
Ssc.29636.1.A1_at	*MAMDC2*	MAM domain containing 2	-2.62	0.0181
Ssc.11787.2.A1_at	*RASSF2*	Ras association (RalGDS/AF-6) domain family member 2	-2.62	0.0020
Ssc.6656.1.A1_at	*FNDC1*	fibronectin type III domain containing 1	-2.61	0.0013
Ssc.4231.1.A1_at	*C1QTNF3*	C1q and tumor necrosis factor related protein 3; alpha-methylacyl-CoA racemase	-2.59	0.0270
Ssc.9598.1.A1_at	*PGM5*	phosphoglucomutase 5	-2.57	0.0116
Ssc.17399.1.S1_at	*KCNMA1*	potassium large conductance calcium-activated channel, subfamily M, alpha member 1	-2.57	0.0047
Ssc.2675.1.S1_at	*TGFB1I1*	transforming growth factor beta 1 induced transcript 1	-2.57	0.0030
Ssc.6343.1.S1_at	*CNTN1*	contactin 1	-2.56	0.0166
Ssc.17375.1.S1_at	*C6orf117*	No Annotatation	-2.56	0.0020
Ssc.19592.3.S1_a_at	*NCS1*	frequenin homolog (Drosophila)	-2.56	0.0013
Ssc.6503.1.A1_at	*Q6ZWD1*	No annotation	-2.56	0.0036
Ssc.7294.1.A1_at	*ECM2*	extracellular matrix protein 2, female organ and adipocyte specific	-2.55	0.0081
Ssc.27703.1.S1_at	*FBN2*	fibrillin 2	-2.54	0.0014
Ssc.29763.1.A1_at	*ENDOU*	26 serine protease	-2.54	0.0007
Ssc.27277.1.S1_at	*TES*	testis derived transcript (3 LIM domains)	-2.52	0.0007
Ssc.6069.1.A1_at	*KCNK9*	potassium channel, subfamily K, member 9	-2.52	0.0086
Ssc.29135.1.S1_at	*CAMK2A*	calcium/calmodulin-dependent protein kinase II alpha	-2.51	0.0036
Ssc.24415.1.A1_at	*NP_115914*	No annotation	-2.50	0.0026
Ssc.16609.1.S1_at	*USP9Y*	ubiquitin specific peptidase 9, Y-linked	-2.49	0.0030
Ssc.21237.1.S1_at	*T2AY*	No annotation	-2.49	0.0053
Ssc.9607.1.A1_at	*RCAN1*	regulator of calcineurin 1	-2.49	0.0013
Ssc.13956.2.S1_a_at	*TPM1*	tropomyosin 1 (alpha)	-2.47	0.0013
Ssc.5047.1.A1_at	*LYVE1*	lymphatic vessel endothelial hyaluronan receptor 1	-2.47	0.0016
Ssc.12900.1.A1_at	*C1QTNF2*	C1q and tumor necrosis factor related protein 2	-2.44	0.0038
Ssc.27341.1.S1_a_at	*SLMAP*	sarcolemma associated protein	-2.44	0.0082
Ssc.9586.2.S1_at	*SDPR*	serum deprivation response (phosphatidylserine binding protein)	-2.44	0.0177
Ssc.4705.2.A1_at	*SMOC2*	SPARC related modular calcium binding 2	-2.44	0.0078
Ssc.16704.1.S1_at	*TACC1*	transforming, acidic coiled-coil containing protein 1	-2.43	0.0016
Ssc.16311.1.A1_a_at	*LHCGR*	luteinizing hormone/choriogonadotropin receptor	-2.43	0.0410
Ssc.25054.1.S1_at	*SLC22A15*	solute carrier family 22, member 15	-2.43	0.0067
Ssc.8072.1.A1_at	*LTBP1*	latent transforming growth factor beta binding protein 1	-2.42	0.0085
Ssc.5956.1.S1_at	*TFPI2*	tissue factor pathway inhibitor 2	-2.41	0.0052
Ssc.1089.1.S1_at	*FABP4*	fatty acid binding protein 4, adipocyte	-2.41	0.0032
Ssc.941.1.S1_at	*ACTA2*	actin, alpha 2, smooth muscle, aorta	-2.41	0.0021
Ssc.6531.1.A1_at	*SPARCL1*	SPARC-like 1 (hevin)	-2.41	0.0120
Ssc.16989.1.A1_at	*CAV1*	caveolin 1, caveolae protein, 22kDa	-2.40	0.0030
Ssc.24909.1.S1_at	*LAMA4*	laminin, alpha 4	-2.39	0.0032
Ssc.15772.1.A1_at	*KCNE2*	potassium voltage-gated channel, Isk-related family, member 2	-2.38	0.0077

Ssc.1808.1.S1_at	*CNNM2*	cyclin M2	-2.38	0.0026
Ssc.7463.1.A1_at	*NP_057165*	No annotation	-2.38	0.0007
Ssc.14279.1.A1_at	*EZH2*	enhancer of zeste homolog 2 (Drosophila)	-2.37	0.0120
Ssc.19883.1.S1_at	*GRIK2*	glutamate receptor, ionotropic, kainate 2	-2.37	0.0014
Ssc.16663.1.S1_at	*ITGA5*	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	-2.37	0.0081
Ssc.9718.1.A1_at	*CITED2*	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	-2.37	0.0065
Ssc.7712.1.A1_at	*TACC1*	transforming, acidic coiled-coil containing protein 1	-2.36	0.0058
Ssc.15710.2.S1_at	*ECHDC1*	enoyl Coenzyme A hydratase domain containing 1	-2.36	0.0032
Ssc.25186.1.A1_at	*ZFPM2*	zinc finger protein, multitype 2	-2.36	0.0130
Ssc.2971.1.S1_at	*TNS1*	tensin 1	-2.36	0.0010
Ssc.5133.1.S1_at	*FXYD1*	FXYD domain containing ion transport regulator 1	-2.36	0.0030
Ssc.8072.2.A1_at	*LTBP1*	latent transforming growth factor beta binding protein 1	-2.33	0.0098
Ssc.8671.1.A1_at	*PRDM5*	PR domain containing 5	-2.33	0.0043
Ssc.11858.1.S1_at	*FMOD*	fibromodulin	-2.33	0.0111
Ssc.12384.1.A1_at	*XK*	X-linked Kx blood group (McLeod syndrome)	-2.32	0.0007
Ssc.7306.1.A1_at	*ITGA9*	integrin, alpha 9	-2.31	0.0317
Ssc.16164.1.S1_at	*MYLK*	myosin light chain kinase	-2.31	0.0032
Ssc.5941.1.S1_a_at	*ACTN1*	actinin, alpha 1	-2.31	0.0014
Ssc.1377.1.A1_at	*ITGA8*	integrin, alpha 8	-2.31	0.0201
Ssc.9450.1.S1_at	*MYLK*	myosin light chain kinase	-2.31	0.0189
Ssc.20515.1.S1_at	*EFEMP1*	EGF-containing fibulin-like extracellular matrix protein 1	-2.30	0.0036
Ssc.27627.1.S1_at	*PPP1R12B*	protein phosphatase 1, regulatory (inhibitor) subunit 12B	-2.29	0.0063
Ssc.17874.1.A1_at	*TOX3*	TOX high mobility group box family member 3	-2.29	0.0083
Ssc.7712.2.S1_at	*TACC1*	transforming, acidic coiled-coil containing protein 1	-2.29	0.0035
Ssc.23368.1.A1_at	*COL4A5*	collagen, type IV, alpha 5	-2.29	0.0027
Ssc.8079.1.S1_at	*NP_061054*	No annotation	-2.29	0.0024
Ssc.1407.1.A1_at	*AKAP2*	A kinase (PRKA) anchor protein 2; paralemmin 2; PALM2-AKAP2 readthrough transcript	-2.28	0.0032
Ssc.6726.2.S1_a_at	*KANK1*	KN motif and ankyrin repeat domains 1; similar to ankyrin repeat domain protein 15 isoform b	-2.27	0.0038
Ssc.4705.1.S1_at	*CNTN2*	contactin 2 (axonal)	-2.27	0.0085
Ssc.9720.1.A1_at	*CYR61*	cysteine-rich, angiogenic inducer, 61	-2.25	0.0043
Ssc.4274.1.S1_at	*LRP1B*	low density lipoprotein-related protein 1B (deleted in tumors)	-2.25	0.0138
Ssc.27341.2.A1_at	*SLMAP*	sarcolemma associated protein	-2.24	0.0111
Ssc.5179.1.A1_at	*CAMKK2*	calcium/calmodulin-dependent protein kinase kinase 2, beta	-2.22	0.0035
Ssc.31124.1.S1_at	*COL14A1*	collagen, type XIV, alpha 1	-2.22	0.0025
Ssc.7106.1.S1_at	*CDO1*	cysteine dioxygenase, type I	-2.21	0.0474
Ssc.8125.1.A1_at	*NPTXR*	neuronal pentraxin receptor	-2.21	0.0103
Ssc.7296.1.A1_at	*PROX1*	prospero homeobox 1	-2.21	0.0047
Ssc.18135.1.S1_at	*PRELP*	proline/arginine-rich end leucine-rich repeat protein	-2.21	0.0030
Ssc.18217.1.S1_at	*RSPO3*	R-spondin 3 homolog (Xenopus laevis)	-2.21	0.0256
Ssc.8325.1.A1_at	*C6orf65*	No Annotatation	-2.20	0.0061
Ssc.23179.1.A1_at	*INPP5A*	inositol polyphosphate-5-phosphatase, 40kDa	-2.19	0.0007
Ssc.9586.1.S1_at	*SDPR*	serum deprivation response (phosphatidylserine binding protein)	-2.19	0.0035
Ssc.21806.1.A1_at	*LHFP*	lipoma HMGIC fusion partner	-2.19	0.0031
Ssc.17313.1.A1_at	*CALD1*	caldesmon 1	-2.19	0.0058
Ssc.19282.1.S1_at	*Q6ZR64*	No annotation	-2.19	0.0007
Ssc.2336.1.A1_at	*NP_055544*	No annotation	-2.19	0.0017
Ssc.2017.1.S1_at	*CSRP1*	cysteine and glycine-rich protein 1	-2.18	0.0011
Ssc.992.1.S1_at	*THBS2*	thrombospondin 2	-2.18	0.0043
Ssc.3189.1.A1_at	*RCAN1*	regulator of calcineurin 1	-2.17	0.0025

Ssc.12936.1.A1_at	*KERA*	keratocan	-2.16	0.0082
Ssc.16231.1.S1_a_at	*IGF1*	insulin-like growth factor 1 (somatomedin C)	-2.16	0.0016
Ssc.3988.1.S1_at	*MAP1A*	microtubule-associated protein 1A	-2.16	0.0019
Ssc.10132.1.S1_at	*FERMT2*	fermitin family homolog 2 (Drosophila)	-2.15	0.0060
Ssc.14062.1.S1_at	*IGFBP6*	insulin-like growth factor binding protein 6	-2.15	0.0034
Ssc.22803.1.S1_at	*C6orf133*	No Annotatation	-2.12	0.0221
Ssc.10303.1.A1_at	*PDGFD*	platelet derived growth factor D	-2.12	0.0130
Ssc.1049.1.S1_at	*COL12A1*	collagen, type XII, alpha 1	-2.11	0.0114
Ssc.18080.1.A1_at	*Q8TAG6*	No annotation	-2.11	0.0356
Ssc.12068.1.A1_at	*COL6A3*	collagen, type VI, alpha 3	-2.11	0.0081
Ssc.6156.1.A1_at	*NBL1*	neuroblastoma, suppression of tumorigenicity 1	-2.11	0.0030
Ssc.26219.1.S1_at	*FXYD6*	FXYD domain containing ion transport regulator 6	-2.10	0.0020
Ssc.83.1.S1_at	*ENPEP*	glutamyl aminopeptidase (aminopeptidase A)	-2.09	0.0477
Ssc.12523.1.A1_at	*CLIC4*	chloride intracellular channel 4	-2.09	0.0155
Ssc.17300.1.S1_at	*TAGLN2*	transgelin 2	-2.09	0.0328
Ssc.11071.2.S1_a_at	*VCL*	vinculin	-2.08	0.0042
Ssc.27409.1.S1_at	*JPH2*	junctophilin 2	-2.08	0.0032
Ssc.12977.1.A1_at	*PDLIM4*	PDZ and LIM domain 4	-2.08	0.0027
Ssc.20464.1.S1_at	*GLIPR1*	GLI pathogenesis-related 1	-2.08	0.0108
Ssc.2491.1.S1_at	*CAMK2G*	calcium/calmodulin-dependent protein kinase II gamma	-2.07	0.0019
Ssc.26335.1.S1_at	*SMTN*	smoothelin	-2.07	0.0249
Ssc.29187.1.A1_at	*GJA1*	gap junction protein, alpha 1, 43kDa	-2.07	0.0392
Ssc.13948.1.S1_at	*GET1*	No annotation	-2.07	0.0017
Ssc.10237.1.S1_at	*HSPB8*	heat shock 22kDa protein 8	-2.07	0.0013
Ssc.9176.1.A1_at	*AEBP2*	AE binding protein 2	-2.06	0.0084
Ssc.5284.1.S1_at	*SLC25A4*	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	-2.06	0.0013
Ssc.17364.1.S1_at	*FAM35A*	family with sequence similarity 35, member A	-2.06	0.0033
Ssc.30886.1.S1_at	*RNGTT*	RNA guanylyltransferase and 5'-phosphatase	-2.06	0.0212
Ssc.3327.1.A1_at	*NP_659493*	No annotation	-2.06	0.0016
Ssc.13641.1.A1_at	*NELL2*	NEL-like 2 (chicken)	-2.05	0.0117
Ssc.6564.1.S1_at	*RASL11A*	RAS-like, family 11, member A	-2.05	0.0405
Ssc.29980.1.S1_at	*DCLK1*	doublecortin-like kinase 1	-2.05	0.0053
Ssc.24603.1.S1_at	*C2orf23*	No Annotatation	-2.05	0.0016
Ssc.10974.1.S1_at	*ALOX15*	arachidonate 15-lipoxygenase	-2.04	0.0183
Ssc.17429.1.S1_at	*PCSK1N*	proprotein convertase subtilisin/kexin type 1 inhibitor	-2.04	0.0054
Ssc.9688.1.S1_at	*NP_079045*	No annotation	-2.03	0.0319
Ssc.19188.2.A1_at	*NP_067651*	No annotation	-2.02	0.0015
Ssc.13345.1.A1_at	*DNAJB4*	DnaJ (Hsp40) homolog, subfamily B, member 4	-2.02	0.0173
Ssc.9299.1.S1_at	*ZEB1*	zinc finger E-box binding homeobox 1	-2.02	0.0041
Ssc.10137.1.A1_at	*LOXL1*	lysyl oxidase-like 1	-2.02	0.0120
Ssc.19579.2.S1_at	*NR2F2*	nuclear receptor subfamily 2, group F, member 2	-2.01	0.0047
Ssc.16671.1.S1_at	*TGFBI*	transforming growth factor, beta-induced, 68kDa	-2.01	0.0136
Ssc.29811.1.A1_at	*AKAP2*	A kinase (PRKA) anchor protein 2; paralemmin 2; PALM2-AKAP2 readthrough transcript	-2.01	0.0101
Ssc.998.1.A1_at	*C1orf24*	No Annotatation	-2.01	0.0336
Ssc.2413.1.S1_at	*SCNN1B*	sodium channel, nonvoltage-gated 1, beta	2.00	0.0089
Ssc.19839.1.S1_at	*PSAT1*	chromosome 8 open reading frame 62	2.01	0.0255
Ssc.493.1.S1_at	*EPB41L4B*	erythrocyte membrane protein band 4.1 like 4B	2.01	0.0065
Ssc.30195.1.A1_at	*NP_005790*	No annotation	2.01	0.0238
Ssc.13115.1.A1_at	*CXADR*	coxsackie virus and adenovirus receptor pseudogene 2	2.01	0.0128
Ssc.19907.1.S1_at	*F3*	coagulation factor III (thromboplastin, tissue factor)	2.02	0.0344

Ssc.30181.1.A1_at	*Q6ZRV2*	No annotation	2.02	0.0082
Ssc.21830.1.A1_at	*GJB3*	gap junction protein, beta 3, 31kDa	2.03	0.0235
Ssc.2159.1.S1_at	*OVOL1*	ovo-like 1(Drosophila)	2.04	0.0032
Ssc.26257.1.S1_at	*MARVELD2*	MARVEL domain containing 2	2.05	0.0306
Ssc.15137.1.A1_at	*C14orf66*	No Annotatation	2.05	0.0058
Ssc.16674.1.S1_at	*CLDN4*	claudin 4	2.05	0.0236
Ssc.17222.1.S1_at	*MUC1*	mucin 1, cell surface associated	2.06	0.0162
Ssc.11074.1.S1_at	*CFD*	complement factor D (adipsin)	2.06	0.0053
Ssc.17029.1.A1_at	*ABO*	ABO blood group (transferase A, alpha 1-3-N-acetylgalactosaminyltransferase	2.06	0.0142
Ssc.1272.1.A1_at	*KEL*	Kell blood group, metallo-endopeptidase	2.06	0.0030
Ssc.24238.1.S1_s_at	*RECQL5*	RecQ protein-like 5	2.06	0.0270
Ssc.4266.1.S1_at	*C20orf42*	No Annotatation	2.07	0.0012
Ssc.3055.1.S1_at	*Q8N412*	No annotation	2.07	0.0217
Ssc.1623.1.S1_at	*DSP*	desmoplakin	2.08	0.0283
Ssc.17490.1.S1_at	*TRDN*	triadin	2.08	0.0496
Ssc.2905.1.S1_at	*SPINT1*	serine peptidase inhibitor, Kunitz type 1	2.08	0.0040
Ssc.4779.1.A1_at	*CPM*	carboxypeptidase M	2.10	0.0257
Ssc.7146.1.A1_at	*ABCA1*	ATP-binding cassette, sub-family A (ABC1), member 1	2.10	0.0261
Ssc.12677.1.S1_at	*ITPR3*	inositol 1,4,5-triphosphate receptor, type 3	2.11	0.0045
Ssc.1510.1.S1_at	*PRSS8*	protease, serine, 8	2.11	0.0106
Ssc.25395.1.S1_at	*PPL*	periplakin	2.13	0.0223
Ssc.5887.1.A1_at	*SLC37A2*	solute carrier family 37 (glycerol-3-phosphate transporter), member 2	2.13	0.0149
Ssc.3610.1.A1_at	*Q96MJ8*	No annotation	2.16	0.0059
Ssc.10743.1.A1_at	*PKP2*	plakophilin 2	2.17	0.0005
Ssc.3651.1.A1_at	*CCT6B*	chaperonin containing TCP1, subunit 6B (zeta 2)	2.19	0.0365
Ssc.13992.1.A1_at	*KLF5*	Kruppel-like factor 5 (intestinal)	2.19	0.0229
Ssc.14108.1.S1_at	*F11R*	F11 receptor	2.20	0.0142
Ssc.7864.1.A1_at	*IL1RAP*	interleukin 1 receptor accessory protein	2.20	0.0172
Ssc.1289.1.S1_at	*HMGA1*	hypothetical LOC100130009	2.23	0.0093
Ssc.1707.1.S1_at	*ARSE*	arylsulfatase E (chondrodysplasia punctata 1)	2.23	0.0013
Ssc.27334.1.S1_at	*NHSL1*	NHS-like 1	2.24	0.0089
Ssc.2444.2.A1_a_at	*LAMA3*	laminin, alpha 3	2.24	0.0102
Ssc.3584.1.S1_at	*DSG2*	desmoglein 2	2.26	0.0045
Ssc.8947.1.S1_at	*BHLHE40*	basic helix-loop-helix family, member e40	2.27	0.0222
Ssc.9063.1.A1_at	*F34A*	No annotation	2.27	0.0123
Ssc.5991.1.A1_at	*KRT18*	keratin 18	2.27	0.0085
Ssc.17615.1.S1_at	*NME7*	non-metastatic cells 7, protein expressed in (nucleoside-diphosphate kinase)	2.29	0.0130
Ssc.14364.1.A1_at	*PPP4R4*	protein phosphatase 4, regulatory subunit 4	2.30	0.0240
Ssc.8861.1.A1_at	*BHLHE40*	basic helix-loop-helix family, member e40	2.30	0.0246
Ssc.24373.1.S1_at	*SLC7A4*	solute carrier family 7 (cationic amino acid transporter, y+ system), member 4	2.31	0.0058
Ssc.2873.1.S1_at	*MA17*	No Annotation	2.32	0.0049
Ssc.3055.2.S1_at	*Q8N412*	No annotation	2.33	0.0270
Ssc.6583.1.S1_at	*CXCL16*	chemokine (C-X-C motif) ligand 16	2.33	0.0053

Ssc.6765.1.S1_at	*SCIN*	scinderin	2.33	0.0171
Ssc.24356.1.S1_at	*EPB41L4B*	erythrocyte membrane protein band 4.1 like 4B	2.34	0.0080
Ssc.22436.1.S1_at	*CYP26B1*	cytochrome P450, family 26, subfamily B, polypeptide 1	2.34	0.0068
Ssc.15588.1.S2_at	*IGFBP3*	insulin-like growth factor binding protein 3	2.36	0.0242
Ssc.11646.2.A1_a_at	*BAIAP2*	BAI1-associated protein 2	2.36	0.0020
Ssc.30283.1.A1_at	*NP_078902*	No annotation	2.40	0.0078
Ssc.23981.1.A1_at	*GALNT12*	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12)	2.41	0.0097
Ssc.16926.1.S1_at	*SLC2A5*	solute carrier family 2 (facilitated glucose/fructose transporter), member 5	2.41	0.0012
Ssc.21691.2.S1_at	*NP_835229*	No annotation	2.42	0.0029
Ssc.314.1.S1_at	*ADM*	adrenomedullin	2.42	0.0075
Ssc.15588.1.S1_at	*IGFBP3*	insulin-like growth factor binding protein 3	2.44	0.0087
Ssc.4135.2.A1_at	*AP1S2*	adaptor-related protein complex 1, sigma 2 subunit pseudogene	2.44	0.0371
Ssc.4190.1.S1_at	*BMP2*	bone morphogenetic protein 2	2.44	0.0013
Ssc.2381.1.A1_at	*S100A9*	S100 calcium binding protein A9	2.46	0.0051
Ssc.18554.1.S1_x_at	*HLA-B*	major histocompatibility complex, class I, C	2.47	0.0033
Ssc.779.1.S1_at	*ELF3*	E74-like factor 3 (ets domain transcription factor, epithelial-specific)	2.47	0.0118
Ssc.6327.1.S1_at	*KRT19*	keratin 19	2.47	0.0036
Ssc.15705.1.A1_at	*KRT8*	keratin 8 pseudogene 9	2.47	0.0024
Ssc.9069.3.S1_at	*OGDHL*	oxoglutarate dehydrogenase-like	2.53	0.0036
Ssc.9069.1.A1_at	*OGDHL*	oxoglutarate dehydrogenase-like	2.53	0.0088
Ssc.24529.1.S1_at	*CRYGS*	crystallin, gamma S	2.53	0.0065
Ssc.25196.1.S1_at	*CKMT1B*	creatine kinase, mitochondrial 1B	2.58	0.0110
Ssc.2143.1.S1_at	*Q8NB62*	No annotation	2.59	0.0052
Ssc.4561.1.S1_at	*FXYD4*	FXYD domain containing ion transport regulator 4	2.60	0.0015
Ssc.14444.3.A1_a_at	*ARG2*	arginase, type II	2.60	0.0055
Ssc.16218.1.S1_at	*ITGA2*	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	2.61	0.0106
Ssc.18435.1.A1_at	*Q8IV31*	No annotation	2.63	0.0032
Ssc.12720.1.S1_at	*TM4SF11*	plasmolipin	2.63	0.0054
Ssc.27507.1.S1_at	-	No annotation	2.64	0.0142
Ssc.3141.1.S1_at	*MAGI1*	membrane associated guanylate kinase, WW and PDZ domain containing 1	2.64	0.0284
Ssc.20578.1.S1_at	*CXCL14*	chemokine (C-X-C motif) ligand 14	2.64	0.0271
Ssc.14414.1.A1_at	*HDC*	histidine decarboxylase	2.65	0.0052
Ssc.29004.1.S1_at	*BMP2*	bone morphogenetic protein 2	2.66	0.0018
Ssc.246.1.S1_at	*ATP1B1*	ATPase, Na+/K+ transporting, beta 1 polypeptide	2.70	0.0054
Ssc.23058.1.A1_at	*C14orf66*	No Annotatation	2.72	0.0061
Ssc.161.1.S1_at	*SLC5A1*	solute carrier family 5 (sodium/glucose cotransporter), member 1	2.72	0.0282
Ssc.28940.1.A1_at	*TRPM6*	transient receptor potential cation channel, subfamily M, member 6	2.73	0.0110
Ssc.9061.1.A1_at	*CST6*	cystatin E/M	2.73	0.0027
Ssc.4033.2.S1_at	*ARSE*	arylsulfatase E (chondrodysplasia punctata 1)	2.82	0.0010
Ssc.24494.1.A1_at	*HDC*	histidine decarboxylase	2.83	0.0046
Ssc.16250.1.S2_at	*IL1RN*	interleukin 1 receptor antagonist	2.86	0.0022
Ssc.21691.3.S1_at	*NP_835229*	No annotation	2.89	0.0043

Ssc.4984.1.S1_at	*CXCL14*	chemokine (C-X-C motif) ligand 14	2.93	0.0307
Ssc.18348.1.S1_at	*IL18RAP*	interleukin 18 receptor accessory protein	2.96	0.0038
Ssc.9434.1.A1_at	*MARK1*	MAP/microtubule affinity-regulating kinase 1	2.98	0.0037
Ssc.12415.1.A1_at	*UNC93A*	unc-93 homolog A (C. elegans)	3.02	0.0131
Ssc.3968.1.S1_at	*WFDC2*	WAP four-disulfide core domain 2	3.06	0.0059
Ssc.5938.1.A1_at	*SDCBP2*	syndecan binding protein (syntenin) 2	3.10	0.0087
Ssc.2165.2.S1_a_at	*SFN*	stratifin	3.18	0.0025
Ssc.4483.1.A1_at	*SLC19A3*	solute carrier family 19, member 3	3.20	0.0277
Ssc.569.1.S1_at	*FXYD3*	FXYD domain containing ion transport regulator 3	3.22	0.0054
Ssc.29575.1.A1_at	*VNN2*	vanin 2	3.26	0.0136
Ssc.7272.1.A1_at	*SERPINB2*	serpin peptidase inhibitor, clade B (ovalbumin), member 2	3.30	0.0040
Ssc.12367.1.A1_at	*MUC4*	mucin 4, cell surface associated	3.34	0.0139
Ssc.3012.1.S1_at	*UPP1*	uridine phosphorylase 1	3.44	0.0165
Ssc.1191.1.S1_at	*FAF1*	Fas (TNFRSF6) associated factor 1	3.54	0.0113
Ssc.23994.1.A1_at	*PTGS2*	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	3.56	0.0030
Ssc.9991.1.S1_at	*PTHLH*	parathyroid hormone-like hormone	3.57	0.0165
Ssc.338.1.S1_at	*MU5A_HUMAN*	No Annotation	3.59	0.0134
Ssc.12431.1.A1_at	*MYO5B*	similar to acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase)	3.60	0.0142
Ssc.760.1.S1_at	*CES3*	carboxylesterase 3	3.68	0.0013
Ssc.7314.1.A1_at	*PTGS2*	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	3.72	0.0211
Ssc.12640.1.A1_at	*PCFB*	No annotation	3.86	0.0083
Ssc.429.1.A1_at	*S100G*	S100 calcium binding protein G	4.10	0.0090

[a]Positive and negative signs indicate whether the level of gene expression is higher in the first or in the second group, respectively.

[b]Adjusted *P-Value* by FDR. Significance Threshold was set at 0.05

3.2.2. Description of a new RT-qPCR methodology for measure miRNA expression

Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers

Balcells, I., Cirera, S. and Busk, P.K.

BMC Biotechnology (Accepted)

Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers

I. Balcells[1]*, S. Cirera[2]* and Peter K Busk[3§]

[1]Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[2]Department of Animal and Veterinary Basic Sciences, University of Copenhagen, Copenhagen, Denmark

[3]Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Lautrupvang 15, 2750 Ballerup, Denmark

*These authors contributed equally to this work

[§]Corresponding author

ABSTRACT

Background

MicroRNAs are important regulators of gene expression at the post-transcriptional level and play an important role in many biological processes. Due to the important biological role it is of great interest to quantitatively determine their expression level in different biological settings.

Results

We describe a PCR method for quantification of microRNAs based on a single reverse transcription reaction for all microRNAs combined with real-time PCR with two, microRNA-specific DNA primers. Primer annealing temperatures were optimized by adding a DNA tail to the primers and could be designed with a success rate of 94 %. The method was able to quantify synthetic templates over eight orders of magnitude and readily discriminated between microRNAs with single nucleotide differences. Importantly, PCR with DNA primers yielded significantly higher amplification efficiencies of biological samples than a similar method based on locked nucleic acids-spiked primers, which is in agreement with the observation that locked nucleic acid interferes with efficient amplification of short templates. The higher amplification efficiency of DNA primers translates into higher sensitivity and precision in microRNA quantification.

Conclusions

MiR-specific quantitative RT-PCR with DNA primers is a highly specific, sensitive and accurate method for microRNA quantification.

BACKGROUND

MicroRNAs (miRNAs) are small non-coding RNAs that are important regulators of biological processes in animals and plants. MiRNAs regulate gene expression at the posttranscriptional level by binding to mRNAs and either inhibit translation or modify the stability of the mRNA. Due to the important biological role of miRNAs it is of great interest to study their expression level in the cells. Furthermore, miRNAs have been associated with cancer and other diseases (1) and miRNA expression can help in the diagnosis and prognostic of human disease (2, 3). The discovery of miRNAs in blood and their surprisingly high stability holds great promise for diagnosis of human disease with miRNAs as biomarkers (4). Several studies have shown that the amount of individual miRNAs in blood is affected by human disease and that the level of specific miRNAs can be used as a diagnostic tool (for examples see (5-9)).

The three methods most frequently used for detection of miRNAs are high-throughput sequencing, microarrays and reverse transcription quantitative PCR (RT qPCR). The latter method is used independently and for validating data obtained from high-throughput sequencing and microarrays. It is challenging to design PCR primers for miRNAs as the typical miRNA is only 22 bases long, which is about the same size as a conventional PCR primer. Several methods have been developed to overcome this problem. Chen and coworkers (10) developed stem–loop RT–PCR where reverse transcription is done at low temperature with a specially designed loop-primer followed by PCR with one specific primer and a universal primer. The PCR product is detected with a TaqMan probe. Although the method requires a specific RT primer for each miRNA, this method can be performed as multiplex so that one RT reaction can be used as template for several qPCR reactions (11). Unfortunately, stem–loop RT–PCR does not allow the user to control the specificity of the reaction by melting curve analysis and the TaqMan probe does not contribute to specificity as the probe binds to the part of the cDNA sequence that originates from the RT primer. Thus, if the RT primer binds to another sequence than the miRNA of interest, this will lead to incorporation of the binding site of the TaqMan

probe and this unspecific amplicon will be indistinguishable from the desired PCR product.

The recently published method based on circularization of the miRNA also depends on a specific primer for reverse transcription (12) and may be difficult to adapt to multiplexing. Furthermore, circularization by RNA ligase is sensitive to sequence bias (13).

Another way to perform miRNA qPCR is to add a poly(A) tail to the miRNA and use a tagged poly(T) primer for reverse transcription (14). Subsequently, PCR is performed with a miRNA-specific primer and a universal primer. This method is very convenient when the amount of sample is limiting, which is often the case for samples such as biopsies and microdissected samples, and when miRNA concentrations are low such as in blood, because it only requires a single RT reaction to generate a template for detection of all miRNAs. However, as only one specific primer is used for PCR there is little degree of freedom in primer design and specificity could be an issue. Especially the discrimination between closely related miRNAs that differ by only one or a few nucleotides can be difficult using this method.

The method called Universal RT microRNA PCR combines the benefits of a universal RT reaction with the specificity of two miRNA-specific PCR primers (15). The PCR product is detected with the intercalating dye SYBR-Green that allows the control of unwanted PCR products by melting curve analysis. The method relies on poly(A) tailing of the miRNAs followed by reverse transcription with a tagged poly(T) primer. PCR is performed with two specific primers that are spiked with Locked Nucleic Acid (LNA) to increase the Tm and the specificity. Although the PCR reactions are specific and discriminate well between closely related miRNAs, they often exhibit a low amplification efficiency which is a common cause of inaccurate quantification. This is in agreement with the observation that sequences containing LNA are poor templates for most DNA polymerases (16).

In the present study we describe that qPCR with two miRNA-specific DNA primers leads to higher amplification efficiency than qPCR with LNA-spiked primers. In addition, this method has all the benefits regarding freedom of primer design and

specificity of the LNA-based method. Optimization of primer Tm and high specificity of the PCR reaction is achieved by adding a tail to each of the PCR primers.

RESULTS

MiR-specific qPCR of miRNAs combines the benefits of a universal RT reaction with the specificity of two miR-specific primers for qPCR (Figure 1). We designed miR-specific DNA primers (Table 1) and tested them at different concentrations in real-time PCR of synthetic miR templates in a background of salmon sperm DNA. A final concentration of 250 nM of each primer was found to be optimal for qPCR (Figure 2A). This primer concentration gave significantly lower Cq values than 125 nM primer whereas 500 nM primer did not reduce the Cq values further.

The PCR reactions were linear over a range of eight log10 of synthetic template (Figure 2B and C), produced one peak in melting curve analysis (Figure 2D) and exhibited a good correlation between Cq and template concentration (Figure 2C).

Amplification of miRNAs from biological samples yielded similar amplification curves as for synthetic templates (Figure 3A) and melting curve analysis indicated the presence of only one amplicon (Figure 3B). In addition, there was a good correlation between Cq and template concentration over four log10 dilutions when biological samples were used ($R2 \geq 0.98$) (Figure 3C).

To test the hypothesis that LNA can inhibit PCR amplification by decreasing the amplification efficiency we compared the efficiency of amplification of 18 miRNAs from porcine uterus with commercially available LNA-spiked primers sets from Exiqon (Denmark) and with DNA primers without LNA. With LNA-spiked primers, amplification efficiencies ranged from 79 to 95 % for 17 of the 18 assays. The last assay (let-7d) had an apparent efficiency of 85 % but more than one peak appeared in the melting curve analysis of the PCR product (data not shown). This indicates that the assay is unspecific and it was excluded from the analysis of assay efficiency (Table 2). Amplification efficiencies with DNA primers ranged from 84 to 102 % (Table 2) and were significantly higher than with LNA-spiked primers (P-value <

0.001). On average, the PCR reactions with DNA primers yielded 5.0 % higher efficiency than LNA-spiked primers corresponding to a 2.4 fold higher sensitivity after 30 cycles of PCR. Melting curve analysis of the let-7d assay with DNA primers only yielded one peak corroborating that this assay was specific (Figure 2D).

The ability of DNA primers to distinguish between miRNAs with a single base difference was tested for three cases where the one base difference was in the part of the miRNA sequence used for forward primer design and two cases where the difference was in the sequence used for reverse primer design (Figure 4A). On average, qPCR of the specific template gave almost 100-fold higher signal than amplification of the template with a single base difference (Figure 4B). For example, amplification of let-7a with the let-7a assay gave a Cq that was 7.6 cycles lower than amplification of the same amount of let-7e with the let-7a assay corresponding to a difference of 170 fold in favor of the intended template compared to the single base mismatch (Figure 4C).

To investigate the effect of different PCR master mixes on the performance of miR-specific qPCR with DNA primers we compared the amplification of synthetic templates with the QuantiFast SYBR Green PCR master mix (Qiagen, Germany) and the Brilliant III Ultra-Fast QPCR Master Mix (Agilent, USA). There was no difference in amplification efficiency (P-value = 0.69) for the five assays tested (let-7d, miR-20a, miR-21, miR-26a and miR-150) between the two master mixes and all the assays gave one peak in melting curve analysis and were comparable over eight log10 of template concentration (Figure 5). The different Tm (peak of the melting curve) in the two master mixes may probably be attributed to different composition of the buffers.

MiR-specific qPCR of let-7a, miR-21, miR-23a and miR-150 with DNA primers on RNA from six different pig tissues showed expression levels from 8 copies per pg total RNA up to almost 2000 copies per pg total RNA (Table 3). Expression of let-7a was remarkably stable with differences below 5 fold between the six tissues. Regardless of the level of expression (Cqs from 16 to 23) and the type of tissue, the assays yielded products with one peak in melting curve analysis as expected for specific PCR

amplifications (data not shown). The same expression profile of the four miRs in the same six samples (P-values > 0.05) was obtained with LNA primers but the Cq values were one cycle higher on average (data not shown).

DISCUSSION

MiR-specific qPCR is a relatively new method that holds great promise. The use of two miR-specific primers makes the method as specific as stem-loop RT-PCR and the reverse transcription is performed with a universal primer compatible with all qPCR primer pairs and is therefore optimal for analysis of small RNA samples and for high-throughput screening (15). Furthermore, detection with intercalating dye allows characterization of the PCR product by melting curve analysis. MiRNA PCR may produce unwanted side products that can only be detected by melting curve analysis.

Commercially available primers for miR-specific qPCR are spiked with LNA (www.exiqon.com). In the present study we found that qPCR reactions with LNA-spiked primers had a tendency to exhibit low amplification efficiencies, which makes accurate quantification more difficult (17). Although several algorithms that account for amplification efficiency are available to calculate the original template concentration from real-time PCR data (18-21) low amplification efficiency is a sign that the amplification reaction is suboptimal and will in all cases lead to lower sensitivity of the PCR reaction (22). However, we found that DNA primers can be successfully used for miR-specific qPCR and that the use of DNA gives significantly higher amplification efficiencies than LNA-spiked primers. Low Tm is often a problem in case of the short primers designed for a miRNA template. This issue can be solved by spiking LNA into the sequence to increase the Tm (23). However, the same can be achieved by adding an artificial sequence to the 5' end of the primer as done for the stem-loop RT-PCR method (10). In the present report we optimized forward primer Tm to 59 °C by adding an artificial sequence at the 5’ end and found that these primers performed well in miR-specific qPCR. The reverse primer for miR-specific PCR is constructed with a short, specific sequence that varies from 4-8 bases at the 3' end followed by a 15 bases long thymidine stretch as in the RT primer and finally, a

5' end tail (tag) that can be varied in length to optimize the Tm (15). Strictly speaking the primer is not specific as only the last 4 – 8 bases in the 3' end are complementary to the miRNA. However, this short sequence combined with the thymidine stretch is sufficient to confer high specificity to the PCR reaction. E.g. templates without a polyA tail or premiRs that extend the miR at the 3' end are not amplified (15). It was reported that it is necessary to spike an LNA into the reverse primer to avoid aberrant amplification products but this effect was only demonstrated for primers with very high Tm (67 – 68 ºC) (15). We found that when the Tm of the reverse primer is optimized to 59 ºC, which is the optimal Tm for the forward primer, the LNA is no longer crucial for successful PCR.

A possible explanation of the lower amplification efficiency with LNA-spiked primers is that for short targets such as miRNAs the primers that are incorporated into the template during amplification will lead to a high proportion of LNA in the template that will decrease the efficiency of subsequent PCR cycles. This possibility is supported by differences between the solution structure of a DNA:LNA helix and the structure of double-stranded DNA (24) and that nucleotide incorporation opposite to an LNA base may be difficult for some polymerases (16). A second possibility is that the LNA-spiked primers may be more prone to form secondary structures that will lower the efficient primer concentration available to hybridize to the template. Stem-loop RT PCR is performed with DNA primers (10) and should therefore have the same efficiency as miR-specific qPCR with DNA primers provided that the detection method does not influence efficiency. Measurement of the efficiency of 87 stem-loop RT PCR assays gave an average efficiency of 94 % ± 0.09 (25). As expected this efficiency is not significantly different from the average efficiency (91 % ± 0.05) for the 18 miR-specific qPCR assays with DNA primers reported in the present study (*P*-value = 0.17, Student's T-test) but it is higher than the average efficiency (85 % ± 0.05) for the 17 miR-specific qPCR assays with LNA-spiked primers reported in the present study (*P*-value = 0.0001, Student's T-test). It therefore seems that DNA primers give higher amplification efficiency of miRNA templates than LNA-spiked primers independently of whether intercalating dye or TaqMan probes are used for detection.

The lower dissociation rates of double-stranded DNA containing LNA bases (26) suggest that PCR with LNA-spiked primers requires longer denaturation times. However, the recommended protocol (www.exiqon.com) has a denaturation time of 30 seconds which should be more than sufficient.

The use of two specific primers for each miRNA allows for design of several different primer sets. E.g, for a 22 bases sequence the forward primer can be from 15-18 bases long and the reverse primer (specific part) can be from 4-8 bases long and the combination of two primers will still cover all of the sequence. This is in contrast to PCR methods with one specific primer, where the primer should always be as long as possible. One significant advantage of this freedom of design is that when discriminating between two miRNAs with a single base mismatch, it is easier to design primers with the 3' end close to the mismatch position, which is optimal for mismatch discrimination (27). In agreement with this, miR-specific qPCR efficiently discriminates between related miRNAs (www.exiqon.com, this study).

Another indication of the robustness of miR-specific qPCR is that the PCR can be performed in different master mixes both with LNA and with DNA primers (this study).

Of the 18 assays designed for the present study, 17 worked well in qPCR, which is a success rate of 94 % for primer design. For the failed primer set the forward and the reverse primers were able to form primer dimers and redesign of the primers solved this problem. By taking primer dimer formation into account it may be possible to reach even higher design success rates for DNA primers. In contrast, the success rate of LNA-spiked primers is 70 % when dimer formation is ignored and 80% when accounting for putative primer dimer formation (15). Although the primer design data set for both DNA and LNA-spiked primers are limited, the difference suggests that DNA primers may be easier to design than LNA-spiked primers in agreement with that the design rules for LNA-spiked primers are complex and slight variations in LNA number, position and sequence context can yield different results (28).

CONCLUSIONS

In conclusion, miR-specific qPCR is a useful method for miRNA detection and the present study demonstrates that the use of DNA primers without LNA gives high PCR efficiencies that allow for precise quantification of the target.

METHODS

Total RNA preparation

Uterus samples from 40 sows at 30-32 days of gestation were immediately snap-frozen in liquid nitrogen and stored at -80ºC until use. Total RNA was extracted with TRIzol® reagent (Invitrogen).

Other pig tissue samples were collected from a 3-months old Danish production pig, except for the ovary sample that was collected from a 6-months old pig. The samples were immediately snap-frozen in liquid nitrogen and stored at -80ºC until use. Total RNA was extracted with TRI Reagent® (Molecular Research Centre, Inc.) following the manufacturer's recommendations.

Uterus samples were obtained from Spanish pigs raised according to the European animal experimentation ethics law approved by the Ethical and Care Committee at IRTA. The rest of the tissues originated from Danish pigs raised under production conditions according to Danish standards for animal husbandry. Since the Danish animals were not subjected to experimental procedures, ethical approval was not required.

RNA quality was examined on an Agilent 2100 Bioanalyzer with the RNA 6000 Nano Kits (Agilent, Germany) or by visual inspection of the 28S/18S ribosomal bands in an agarose gel. RNA quantity was measured on a Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA).

cDNA synthesis

Total RNA was used for cDNA synthesis essentially as described (15). Briefly, 100 ng of RNA in a final volume of 10 µl including 1 µl of 10x poly(A) polymerase buffer, 0.1 mM of ATP, 1 µM of RT-primer, 0.1 mM of each deoxynucleotide (dATP, dCTP, dGTP and dTTP), 100 units of MuLV reverse transcriptase (New England Biolabs, USA) and 1 unit of poly(A) polymerase (New England Biolabs, USA) was incubated at 42 ºC for 1 hour followed by enzyme inactivation at 95 ºC for 5 minutes. The sequence of the RT-primer was 5'-CAGGTCCAGTTTTTTTTTTTTTTVN, where V is A, C and G and N is A, C, G and T. The primer was purchased from TAG Copenhagen (Denmark).

For the microRNA LNA™ PCR kit from Exiqon (Denmark) cDNA synthesis was done according to the manufacturer's instructions.

Design of PCR primers and synthetic templates

All DNA PCR primers were designed according the design rules as previously described (15) except that no LNAs were spiked into the primers. Instead, Tm was optimized to 59 °C by adjusting the tail length of the primers. Tm was calculated according to the nearest-neighbor model (29). Special attention was taken to design the 3' end of the primers according to the following rules:

1. Discard all A's from the 3' end of the miRNA sequence.
2. Choose the longest possible forward primer (12 to 18 bases long) that leaves at least four bases at the 3' end of the miRNA for design of the reverse primer.
3. If possible, the last five bases at the 3' end of the forward primer should include 2-3 A or T residues.
4. If possible, the three last bases at the 3' end of the forward primer should include 1-2 A or T residues.
5. If possible, the two last bases at the 3' end of the forward primer should include 1 A or T residue.
6. If the Tm of the forward primer is below 59 °C add the following bases: G, A, C, G, C at the 5' end one at a time and calculate the Tm. Choose the shortest

of these primers that has a Tm = 59 °C. (E.g. longest possible primer is: $CGCAGN_{18}$, where N_{18} are 18 miR-specific bases and CGCAG is a tail sequence that is not complementary to the miR).

7. If the Tm of the forward primer is above 59 °C remove bases from the 5' end one at a time and calculate the Tm. Choose the longest of these primers that has a Tm = 59 °C.
8. Choose the longest possible reverse primer (4 to 8 bases long) that is not complementary to the 3' end of the forward primer.
9. Choose the reverse primer with the best 3' end according to steps 3-5.
10. Add 15 T's at the 5' end of the reverse primer.
11. If the Tm of the reverse primer is below 59 °C add the following bases at the 5' end one at a time and calculate the Tm: G, A, C, C, T, G, G, A, C. Choose the shortest of these primers that has a Tm = 59 °C. (E.g. longest possible primer is : $CAGGTCCAGT_{15}N_8$, where N_8 are 8 miR-specific bases, T_{15} are 15 T's and CAGGTCCAG is a tail sequence complementary to the tail of the RT primer).

Synthetic templates were DNA oligonucleotides complementary to the mature sequence of the miRNAs including the RT primer sequence that is incorporated during cDNA synthesis. Sequences of primers and templates are given in Table 1. Oligonucleotides were purchased from TAG Copenhagen (Denmark) and Sigma (UK). Primers spiked with LNA were microRNA LNA™ PCR primer sets designed by Exiqon (Denmark).

Quantitative PCR

Quantitative PCR of biological samples was done in 10 µl total volume with 1 µl of cDNA diluted 8-10 times, 5 µl of 2x QuantiFast SYBR Green PCR master mix (Qiagen, Germany), 250nM of each primer (Table 1) or 2 µl microRNA LNA™ primer sets (Exiqon, Denmark). Standard curves with 10-fold dilutions (made with a pool of equal amounts of cDNA from the 40 uterus samples) were made for all assays to calculate qPCR efficiency.

The same PCR conditions were used for synthetic templates except that 1µl of synthetic template in 2 ng/µl salmon sperm DNA (Sigma, USA) in TE was used instead

of cDNA. 2x Brilliant III Ultra-Fast QPCR Master Mix (Agilent, USA) was used instead of QuantiFast where indicated.

Cycling conditions were 95ºC for 5-10 min followed by 40 cycles of 95ºC for 10-30 sec and 60ºC 30-60 sec. A melting curve analysis (60ºC to 99ºC) was performed after the thermal profile to ensure specificity in the amplification.

QPCR of biological samples was performed on a MX3000P machine (Stratagene, USA) and reactions containing synthetic templates were performed on a Rotorcycler (Qiagen, Germany). Primers spiked with LNA were microRNA LNA™ PCR primer sets designed by Exiqon (Denmark).

qPCR data analysis

Quantification was based on determination of the quantification cycle (Cq) and PCR efficiency was calculated from the log-linear portion of the standard curves (17).

Comparison of the efficiency of qPCR with LNA-spiked and DNA primers was done by two-sided Student's T-test for paired samples. Significance threshold was set at *P-value* < 0.05.

COMPETING INTERESTS

PKB is designated as inventor of miR-specific qPCR in a patent filed by Exiqon A/S. All commercial rights to method described in the patent belong to Exiqon A/S. None of the authors have any economical interest in this company.

AUTHORS' CONTRIBUTIONS

PKB designed all oligonucleotides and performed and analyzed all experiments with synthetic templates. IB and SC collected biological samples, purified RNA and performed and analyzed qPCR experiments with these samples. The manuscript was written by the authors from a draft by PKB. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors thank Agnieszka Podolska and Mette Lange for critical comments on the manuscript. This work was supported by the Projects AGL2007-66371-C02-01 and AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 Program (CSD2007-00036) from Ministerio de Ciencia e Innovación. IB is recipient of PIF PhD fellowship from Universitat Autònoma de Barcelona.

REFERENCES

1. Schetter AJ, Heegaard NHH, Harris CC: Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. Carcinogenesis 2010, 31:37-49.

2. Fabbri M: miRNAs as molecular biomarkers of cancer. Expert Rev Mol Diagn 2010, 10:435-444.

3. Ferracin M, Veronese A, Negrini M: Micromarkers: miRNAs in cancer diagnosis and prognosis. Expert Rev Mol Diagn 2010, 10:297-308.

4. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, Vessella RL, Nelson PS, Martin DB, Tewari M: Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U.S.A 2008, 105:10513-10518.

5. Heneghan HM, Miller N, Lowery AJ, Sweeney KJ, Newell J, Kerin MJ: Circulating microRNAs as novel minimally invasive biomarkers for breast cancer. Ann Surg 2010, 251:499-505.

6. Fichtlscherer S, De Rosa S, Fox H, Schwietz T, Fischer A, Liebetrau C, Weber M, Hamm CW, Röxe T, Müller-Ardogan M, Bonauer A, Zeiher AM, Dimmeler S: Circulating microRNAs in patients with coronary artery disease. Circ Res 2010, 107:677-684.

7. Zhang Y, Jia Y, Zheng R, Guo Y, Wang Y, Guo H, Fei M, Sun S: Plasma MicroRNA-122 as a Biomarker for Viral-, Alcohol-, and Chemical-Related Hepatic Diseases. Clin Chem 2010, 56:1830-1838.

8. Vasilescu C, Rossi S, Shimizu M, Tudor S, Veronese A, Ferracin M, Nicoloso MS, Barbarotto E, Popa M, Stanciulea O, Fernandez MH, Tulbure D, Bueso-Ramos CE,

Negrini M, Calin GA: MicroRNA fingerprints identify miR-150 as a plasma prognostic marker in patients with sepsis. PLoS ONE 2009, 4:e7405.

9. Liu C, Kao S, Tu H, Tsai M, Chang K, Lin S: Increase of microRNA miR-31 level in plasma could be a potential marker of oral cancer. Oral Dis 2010, 16:360-364.

10. Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ: Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 2005, 33:e179.

11. Mestdagh P, Feys T, Bernard N, Guenther S, Chen C, Speleman F, Vandesompele J: High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. Nucleic Acids Res 2008, 36:e143.

12. Kumar P, Johnston BH, Kazakov SA: miR-ID: A novel, circularization-based platform for detection of microRNAs. RNA 2011, 17:365-380.

13. Wang H, Ach RA, Curry B: Direct and sensitive miRNA profiling from low-input total RNA. RNA 2007, 13:151-159.

14. Shi R, Chiang VL: Facile means for quantifying microRNA expression by real-time PCR. BioTechniques 2005, 39:519-525.

15. Busk PK: Method for Quantification of Small RNA Species. 2010. WO/2010/085966.

16. Veedu RN, Vester B, Wengel J: Enzymatic incorporation of LNA nucleotides into DNA strands. Chembiochem 2007, 8:490-492.

17. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009, 55:611-622.

18. Pfaffl MW, Horgan GW, Dempfle L: Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 2002, 30:e36.

19. Tichopad A, Dilger M, Schwarz G, Pfaffl MW: Standardized determination of real-time PCR efficiency from a single reaction set-up. Nucleic Acids Res 2003, 31:e122.

20. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM: Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res 2009, 37:e45.

21. Rutledge RG, Stewart D: A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR. BMC Biotechnol 2008, 8:47.

22. Bustin SA: A-Z of Quantitative PCR (IUL Biotechnology, No. 5). International University Line; 2004.

23. Raymond CK, Roberts BS, Garrett-Engele P, Lim LP, Johnson JM: Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs. RNA 2005, 11:1737-1744.

24. Nielsen KE, Singh SK, Wengel J, Jacobsen JP: Solution structure of an LNA hybridized to DNA: NMR study of the d(CT(L)GCT(L)T(L)CT(L)GC):d(GCAGAAGCAG) duplex containing four locked nucleotides. Bioconjug Chem 2000, 11:228-238.

25. Chen Y, Gelfond JAL, McManus LM, Shireman PK: Reproducibility of quantitative RT-PCR array in miRNA expression profiling and comparison with microarray analysis. BMC Genomics 2009, 10:407.

26. Arora A, Kaur H, Wengel J, Maiti S: Effect of locked nucleic acid (LNA) modification on hybridization kinetics of DNA duplex. Nucleic Acids Symp Ser 2008, 52:417-418.

27. Sambrook J, Russell DW: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press; 2001.

28. Latorra D, Arar K, Hurley JM: Design considerations and effects of LNA in PCR primers. Mol Cell Probes 2003, 17:253-259.

29. SantaLucia J: A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U.S.A 1998, 95:1460-1465.

.

TABLES

Table 1. MiRNAs, PCR primers and synthetic templates

miRNA	Sequence	Forward primer	Reverse primer	Synthetic template
let-7a	UGAGGUAGUAGGUUGUAUAGUU	GCAGTGAGGTAGTAGGTTGT	GGTCCAGTTTTTTTTTTTTTTTTAACTATAC	CAGGTCCAGTTTTTTTTTTTTTTTAACTATACAACCTACTACCTCA
let-7d	AGAGGUAGUAGGUUGCAUAGUU	AGAGAGGTAGTAGGTTGCAT	AGGTCCAGTTTTTTTTTTTTTTTAACT	CAGGTCCAGTTTTTTTTTTTTTTTAACTATGCAACCTACTACCTCT
miR-20a	UAAAGUGCUUAUAGUGCAGGUAG	ACAGTAAAGTGCTTATAGTGCA	GTCCAGTTTTTTTTTTTTTTTCTACCT	CAGGTCCAGTTTTTTTTTTTTTTTCTACCTGCACTATAAGCACTTTA
miR-21	UAGCUUAUCAGACUGAUGUUGA	TCAGTAGCTTATCAGACTGATG	CGTCCAGTTTTTTTTTTTTTTTCAAC	CAGGTCCAGTTTTTTTTTTTTTTTCAACATCAGTCTGATAAGCTA
miR-23a	AUCACAUUGCCAGGGAUUUCCA	CATCACATTGCCAGGGAT	CGTCCAGTTTTTTTTTTTTTTTGGAA	CAGGTCCAGTTTTTTTTTTTTTTTGGAAATCCCTGGCAATGTGAT
miR-23b	AUCACAUUGCCAGGGAUUACCAC	same as for miR-23a	TCCAGTTTTTTTTTTTTTTTGTGGTA	CAGGTCCAGTTTTTTTTTTTTTTTGTGGTAATCCCTGGCAATGTGAT
miR-25	CAUUGCACUUGUCUCGGUCUGA	CATTGCACTTGTCTCGGT	GGTCCAGTTTTTTTTTTTTTTTCAGA	
miR-26a	UUCAAGUAAUCCAGGAUAGGCU	CGAGTTCAAGTAATCCAGGA	CCAGTTTTTTTTTTTTTTTAGCCTATC	CAGGTCCAGTTTTTTTTTTTTTTTAGCCTATCCTGGATTACTTGAA
miR-27a	UUCACAGUGGCUAAGUUCCGC	CAGTTCACAGTGGCTAAGA	CAGTTTTTTTTTTTTTTTGCGGAA	CAGGTCCAGTTTTTTTTTTTTTTTGCGGAACTTAGCCACTGTGAA
miR-101a	UACAGUACUGUGAUAACUGAA	CGCAGTACAGTACTGTGATAAC	AGGTCCAGTTTTTTTTTTTTTTTCAG	CAGGTCCAGTTTTTTTTTTTTTTTCAGTTATCACAGTACTGTA
miR-103	AGCAGCAUUGUACAGGGCUAUGA	AGAGCAGCATTGTACAGG	GGTCCAGTTTTTTTTTTTTTTTCATAG	
miR-122	UGGAGUGUGACAAUGGUGUUUGU	ACAGTGGAGTGTGACAATG	TCCAGTTTTTTTTTTTTTTTCAAACAC	CAGGTCCAGTTTTTTTTTTTTTTTACAAACACCATTGTCACACTCCA
miR-125b	UCCCUGAGACCCUAACUUGUGA	CAGTCCCTGAGACCCTA	GTCCAGTTTTTTTTTTTTTTTCACAA	CAGGTCCAGTTTTTTTTTTTTTTTCACAAGTTAGGGTCTCAGGGA
miR-139b-5p	UCUACAGUGCACGUGUCUCCAGU	TCTACAGTGCACGTGTCT	GTCCAGTTTTTTTTTTTTTTTACTGGA	CAGGTCCAGTTTTTTTTTTTTTTTACTGGAGACACGTGCACTGTAGA
miR-150	UCUCCCAACCCUUGUACCAGUG	GTCTCCCAACCCTTGTAC	GTCCAGTTTTTTTTTTTTTTTCACTG	CAGGTCCAGTTTTTTTTTTTTTTTCACTGGTACAAGGGTTGGGAGA
miR-199b-3p	UACAGUAGUCUGCACAUUGGUU	CAGTACAGTAGTCTGCACAT	GTCCAGTTTTTTTTTTTTTTTAACCAA	CAGGTCCAGTTTTTTTTTTTTTTTAACCAATGTGCAGACTACTGTA
miR-200b	UAAUACUGCCUGGUAAUGAUGA	ACAGTAATACTGCCTGGTAATG	GGTCCAGTTTTTTTTTTTTTTTCATC	CAGGTCCAGTTTTTTTTTTTTTTTCATCATTACCAGGCAGTATTA
miR-200c	UAAUACUGCCGGGUAAUGAUGGA	AGTAATACTGCCGGGTAATG	GTCCAGTTTTTTTTTTTTTTTCCATC	CAGGTCCAGTTTTTTTTTTTTTTTCCATCATTACCCGGCAGTATTA

Table 2. Efficiency of miR-specific qPCR assays with LNA-spiked and DNA primers on pig uterus total RNA.

Target	Efficiency LNA primers	Efficiency DNA primers	Difference
let-7a	82%	89%	6.9%
miR-101a	85%	90%	4.9%
miR-103	93%	94%	1.6%
miR-122	95%	95%	-0.1%
miR-125b	89%	94%	4.5%
miR-139b-5p	79%	86%	6.4%
miR-150	84%	97%	12.6%
miR-199b-3p	80%	87%	7.1%
miR-20a	88%	86%	-2.0%
miR-200b	80%	94%	13.6%
miR-200c	83%	84%	0.2%
miR-21	91%	92%	1.1%
miR-23a	79%	93%	14.1%
miR-23b	81%	87%	6.2%
miR-25	84%	91%	6.7%
miR-26a	88%	96%	8.3%
miR-27a	86%	85%	-1.1%
let-7d	not specific	102%	
Average	**85%**	**90%**	**5.4%**

Table 3. Expression profiling of four miRNAs in pig tissues measured by miR-specific qPCR with DNA primers. Expression levels are given as miRNA copy numbers per pg of total RNA.

miRNA	brain	heart	liver	lung	thymus	ovary	Cq(min)	Cq(max)
let-7a	120	87	27	120	34	98	16.2	18.8
miR-21	88	190	36	900	340	1800	15.9	20.7
miR-23a	15	42	8	100	11	33	16.2	20.4
miR-150	39	22	19	140	270	21	18.6	23.4

FIGURES

Figure 1. Flow scheme of miR-specific qPCR. 1. Start with purified RNA containing miRNA. **2.** Add poly(A) tail with poly(A) polymerase (PAP). **3.** cDNA synthesis with reverse transcriptase (RTase) and an anchored poly(T) primer with a 5' tag. **4.** PCR with two tagged primers.

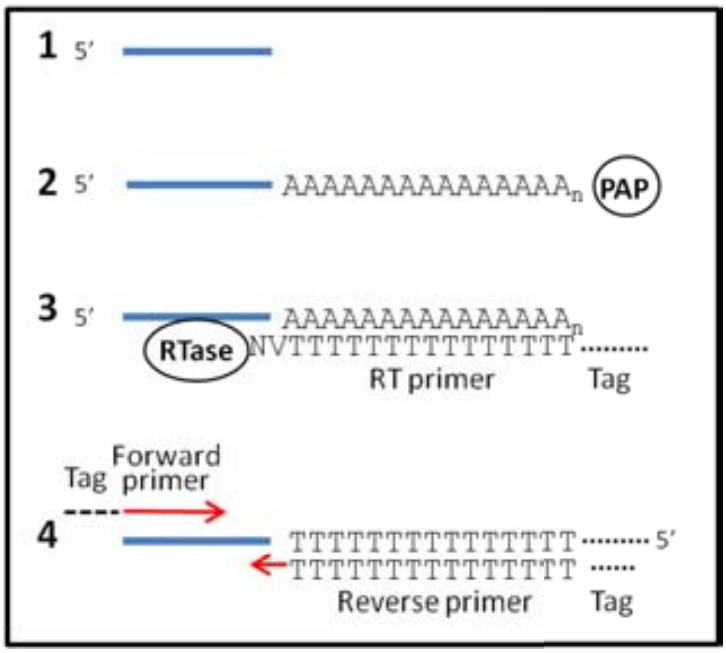


Figure 2. MiR-specific qPCR on synthetic templates with DNA primers. A The effect of primer concentration on Cq value of *ssc-let-7d* and *ssc-miR-26a* miR-specific qPCR assays. Real-time PCR assays were performed in parallel at three different concentrations (125, 250 and 500 nM) of the forward and of the reverse primers. **B** Amplification curves of an eight $\log_{10}$ dilution series of a synthetic *ssc-let-7d* template in the *ssc-let-7d* miR-specific qPCR assays. All samples contained a final concentration of 0.2 ng/µl salmon sperm DNA. **C** Extrapolation of Cq as function of the $\log_{10}$ of the number of templates for the same experiment as in B was a straight line (R^2 = 0.9993) with slope of -3.341 (PCR efficiency = 99 %) over eight $\log_{10}$ dilution of the template. **D** Melting curve analysis of the same experiment. No template control is labeled ntc. Melting curve analysis was performed from 60 ºC to 99 ºC.

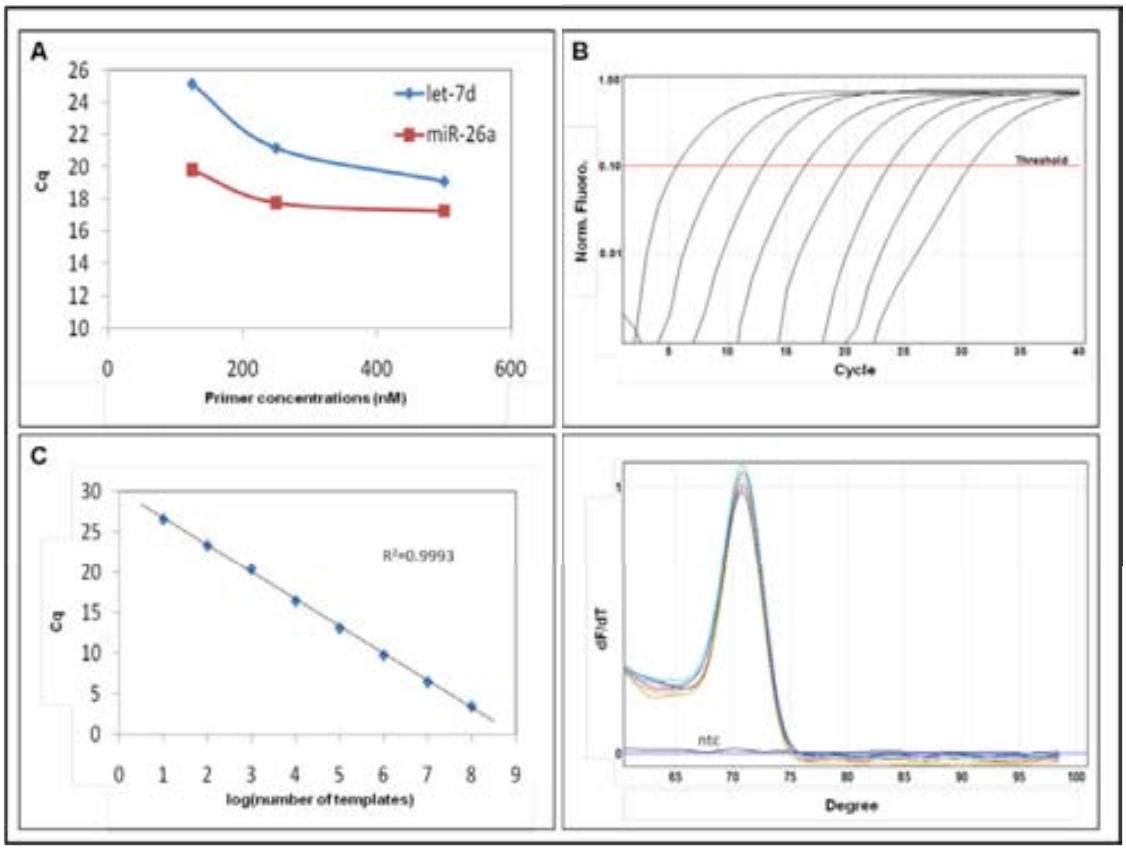


Figure 3. MiR-specific qPCR on biological samples with DNA primers. A Amplification curves of 40 uterus samples with the *ssc-miR-150* miR-specific qPCR assay. **B** Melting curve analysis of the same experiment. Melting curve analysis was performed from 55 ºC to 95 ºC. C Extrapolation of Cq as function of the $\log_{10}$ of the number of templates for the same experiment as in A was a straight line (R^2 = 1.0) with a slope of -3.406 (PCR efficiency = 97 %) over 4 $\log_{10}$ dilution of a pool that includes all samples included in the study.

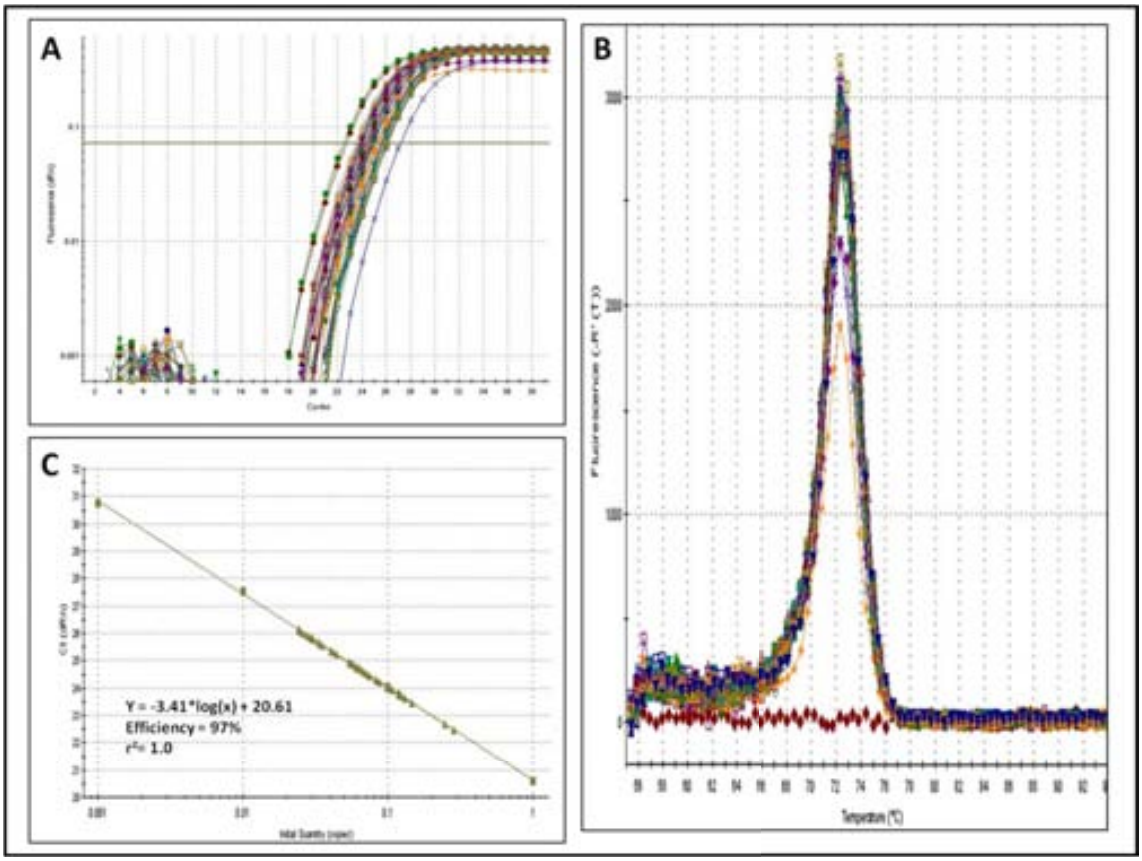
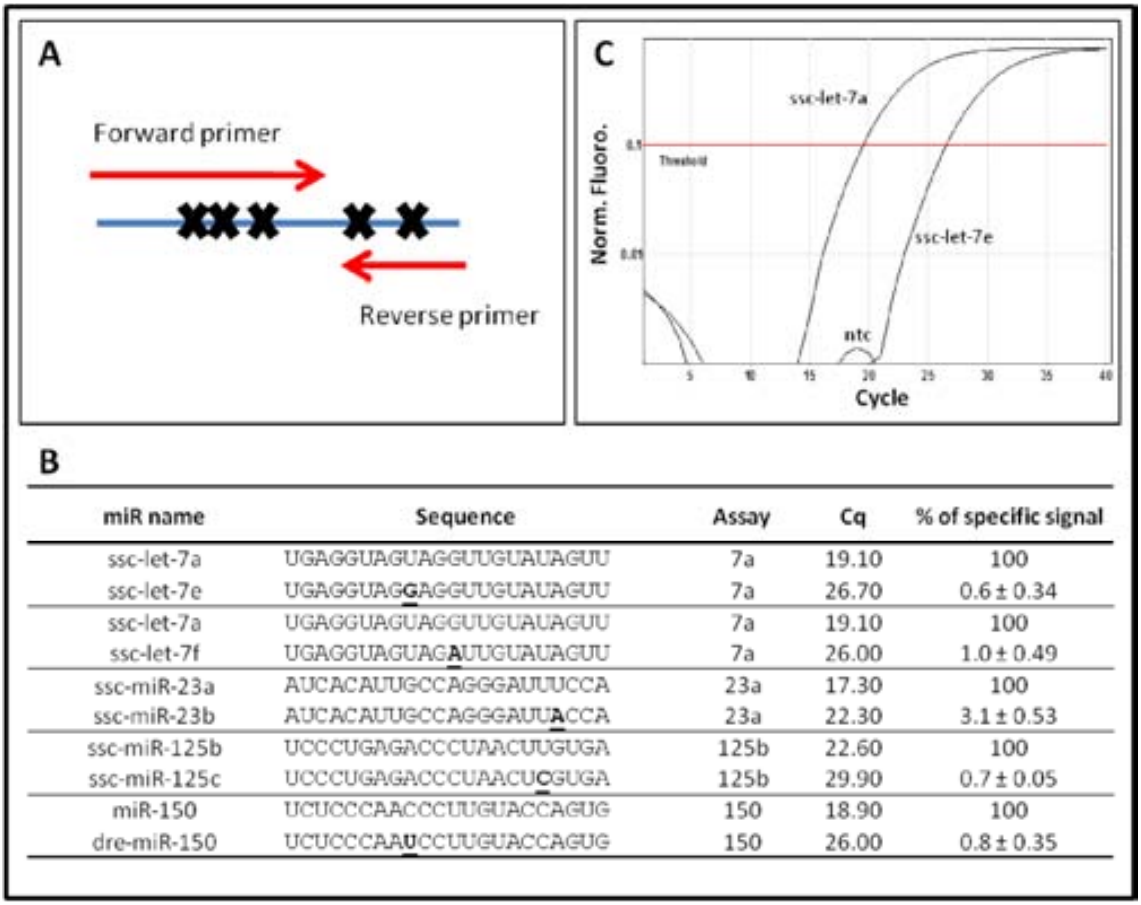


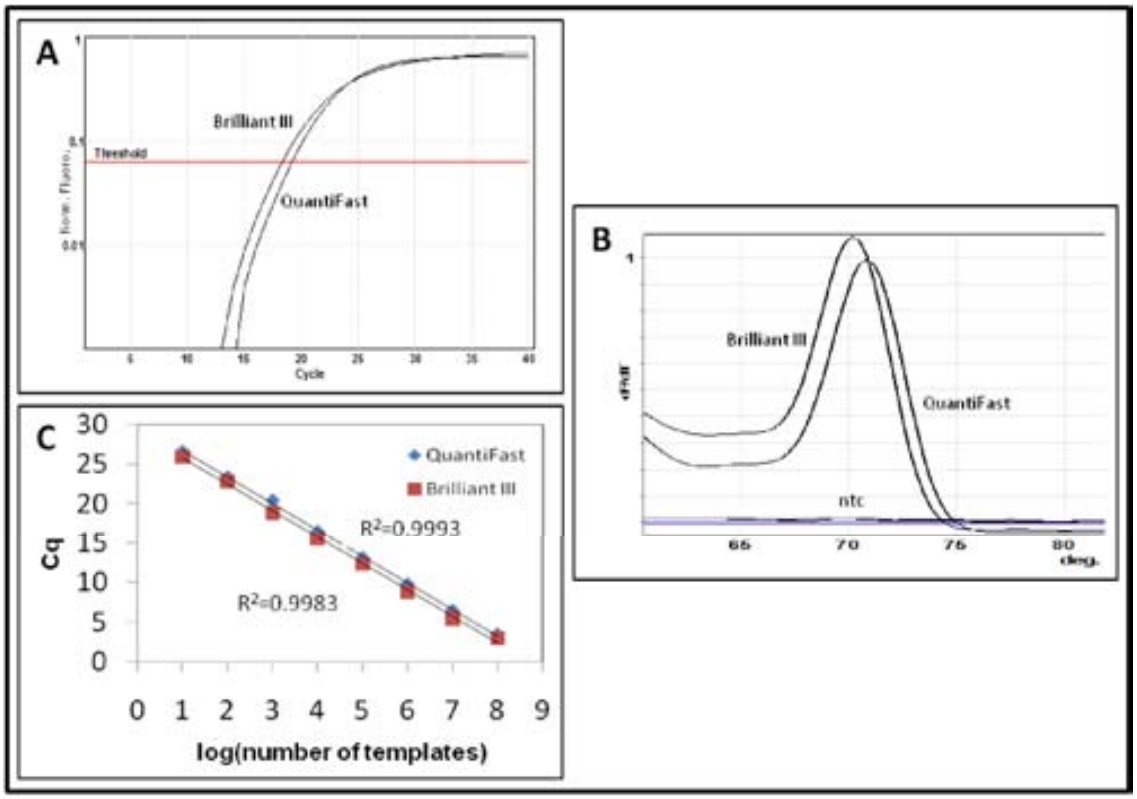
Figure 4. Discrimination between miRNAs with single nucleotide differences. A Position of the single nucleotide mismatches relative to the PCR primers for the *ssc-let-7a, ssc-miR-23a, ssc-miR-125b* and *ssc-miR-150* qPCR assays. The *ssc-miR-23b* sequence used for mismatch discrimination was taken from miRBase and is different from the *ssc-miR-23b* sequence found in uterus and used for designing the *ssc-miR-23b* qPCR primers (Table 1). **B** Discrimination between closely related miRNA templates for miR-specific qPCR assays with DNA primers. Mismatches in the miRNA compared to the PCR primers are underlined. The data represents the results of three to four measurements. **C** Amplification curves of *ssc-let-7a* and *ssc-let-7e* synthetic template in the *ssc-let-7a* miR-specific qPCR assays. All samples including the no template control (ntc) contained a final concentration of 0.2 ng/µl salmon sperm DNA.



B

miR name	Sequence	Assay	Cq	% of specific signal
ssc-let-7a	UGAGGUAGUAGGUUGUAUAGUU	7a	19.10	100
ssc-let-7e	UGAGGUAGGAGGUUGUAUAGUU	7a	26.70	0.6 ± 0.34
ssc-let-7a	UGAGGUAGUAGGUUGUAUAGUU	7a	19.10	100
ssc-let-7f	UGAGGUAGUAGAUUGUAUAGUU	7a	26.00	1.0 ± 0.49
ssc-miR-23a	AUCACAUUGCCAGGGAUUUCCA	23a	17.30	100
ssc-miR-23b	AUCACAUUGCCAGGGAUUACCA	23a	22.30	3.1 ± 0.53
ssc-miR-125b	UCCCUGAGACCCUAACUUGUGA	125b	22.60	100
ssc-miR-125c	UCCCUGAGACCCUAACUCGUGA	125b	29.90	0.7 ± 0.05
miR-150	UCUCCCAACCCUUGUACCAGUG	150	18.90	100
dre-miR-150	UCUCCCAAUCCUUGUACCAGUG	150	26.00	0.8 ± 0.35

Figure 5. MiR-specific qPCR in different qPCR master mixes. A Comparison of amplification curves of a synthetic *ssc-let-7d* template in the *ssc-let-7d* miR-specific qPCR assay in QuantiFast and in Brilliant III qPCR Master mixes. **B** Melting curve analysis of the same experiment. No template control is labeled ntc. Melting curve analysis was performed from 60 ºC to 99 ºC. No change in fluorescence (dF/dT = 0) was observed above 80 ºC and this part of the curves was omitted from the figure. **C** Extrapolation of Cq as function of the $\log_{10}$ of the number of templates for the same experiment as in A was a straight line (R^2 indicated on figure) and for both master mixes the PCR efficiency was 99 % as calculated from the slope of the regression line



3.2.3. Analysis of the microRNAome in the ovary of pregnant IbxMe F_2 sows regarding prolificacy levels.

Differential microRNA expression in the ovaries of high and low prolificacy pregnant sows

Balcells, I., Timoneda, O., Cirera, S., Søkilde, R., Litman, T., Fernández-Rodríguez, A., Noguera, J.L., Sánchez, A., and Tomás, A.

Manuscript in preparation

Differential microRNA expression in the ovaries of high and low prolificacy pregnant sows

Balcells, I.[1,*], Timoneda, O., Cirera, S.[2], Søkilde, R.[3], Litman, T.[4], Fernández-Rodríguez, A.[5], Noguera, J.L.[6], Sánchez, A.[1] and Tomás, A[7].

[1]Departament de Genètica Animal, Centre de Recerca en AgriGenòmica (CRAG), Universitat Autònoma de Barcelona, Bellaterra, Spain

[2]Department of Basic Animal and Veterinary Sciences, Division of Genetics and Bioinformatics, University of Copenhagen, Copenhagen, Denmark

[3]Department of Oncology, University of Lund, Lund, Sweden

[4]Molecular Biomedicine, LEO Pharma A/S, Ballerup, Denmark

[5]Departamento de Mejora Genética Animal, SGIT-INIA, Madrid, Spain

[6]Genètica i Millora Animal, IRTA, Lleida, Spain

[7]Program Infection and Immunity, FISIB, Bunyola, Spain

[*]Correspondence author

ABSTRACT

Background

Reproductive traits, such as prolificacy, are of great economical interest in the pig industry. A good gestation development is assured when there is an accurate temporal and spatial regulation of gene expression in reproductive tissues. microRNAs (miRNAs) are post-transcriptional regulators of gene expression which have been associated to reproductive processes. In the ovary, miRNAs are important during ovulation and corpus luteum development. However, the role of ovarian miRNAs during gestation has not been investigated although ovaries are essential during gestation. Thus, it has been suggested that ovarian miRNAs could participate during the gestation process and, moreover, they could influence the prolificacy levels in sows by regulating gene expression. Overall, it would provide a better understanding of the molecular mechanisms that influence reproductive traits which could help in developing new genetic tools useful for improving the efficiency in pig production.

Results

The miRNA expression profile was compared in the ovaries of pregnant Iberian x Meishan F2 sows displaying extreme phenotypes regarding prolificacy levels defined as the number of embryos (NE) attached to the uterus at 30-32 days of gestation. Differences in ovarian expression for miR-146a-5p and miR-142-3p were identified between high (NE ≥ 13) and low (NE ≤ 11) prolificacy sows. The targets for prolificacy-associated miRNAs were predicted and functionally annotated to further investigate the biological role of these miRNAs. miR-146a-5p is involved mainly in immune system response and also in metabolic and developmental processes. miR-142-3p, on the other hand, participates in the regulation of transcription, intracellular and protein transport, cell communication, regulation of cellular and metabolic processes and cell cycle. The expression of prolificacy-associated miRNAs is negatively correlated with the expression of their mRNA targets: PPM1K, TLR1 and CPEB2 which have been described in a previous study by our group as differentially expressed in the ovaries of pregnant sows regarding the prolificacy levels.

Conclusions

miR-146a-5p and miR-142-3p have been associated with prolificacy levels in ovaries of pregnant sows. Both miRNAs are involved in the immune system and in cellular maintenance which play key roles during reproduction processes. It seems that upregulation of miR-146a-5p and miR-142-3p could help in the establishment of a uterine environment favourable for embryonic development.

BACKGROUND

Reproduction traits, such as fertility and prolificacy, are of great interest to the pig industry. Improving the sow productivity, achieved by increasing the litter size, can have a great economic impact. In addition to external factors like feed, nutrition, husbandry and management practices, genetic factors have a large influence on litter size [1]. A better understanding of the genetic architecture underlying litter size will increase our knowledge of porcine reproduction which could be used to increase the efficiency of pig productivity.

In pigs, prenatal losses during pregnancy have an essential impact on litter size [1]. Importantly, the foetal survival rate from the 31st day of gestation will be an indication of the final litter size [2]. There are two critical phases in the prenatal losses; embryonic or foetal losses depending on the developmental stage of the conceptus. The first phase includes the implantation process between days 12 - 18 of pregnancy where the embryonic losses can reach up to 40 % with an average of 20 – 30% [3]. The second phase occurs during the foetal growing phase, when the embryo is already attached to the endometrium, from day 30 of pregnancy. During this period, foetal losses range between 5% and 50% with an average of 10 - 20% [1].

An accurate temporal and spatial pattern of gene expression during pregnancy is necessary in all reproductive tissues to ensure an optimal gestational process and embryonic development [4]. Most of the studies aimed at determining the gene expression profile in reproductive tissues during pregnancy have been focused mainly on the endometrium, where the gestation process takes place, while the ovary has been commonly neglected [5-7]. Nevertheless, the ovary has an important role during

gestation by regulating the expression of crucial hormones for embryo survival and development [8]. In fact, progesterone which is synthesised by the corpus luteum of the ovary at the early stages of gestation is the responsible for inducing essential changes that lead to a temporary state of uterine receptivity for conceptus implantation [9].

Gene expression is regulated by many epigenetics mechanisms such as DNA methylation, histone modifications, chromatin-remodelling and non-coding RNA (ncRNA) mediated gene-silencing, among others [10]. microRNAs (miRNAs) are a relatively new class of small ncRNAs that regulate gene expression at post-transcriptional level in a wide range of biological processes including reproduction [11-13]. Studies in humans and mice have shown that miRNAs play a key role in reproductive tissues. In the ovary, miRNAs influence ovulation and corpus luteum development [14] whereas in the uterus, they are important at gestation, during the implantation and placenta development [15-17]. Furthermore, miRNAs have been shown to be regulated by steroid hormones which are essential for female reproduction [14, 18]. In humans, dysregulation of miRNAs expression has been associated with ovarian, endometrial and cervical cancer as well as with other important reproductive diseases like premature ovarian failure (POF), endometriosis, uterine fibroids and preeclampsia [11, 19-21]. In mice, knockout models for Dicer, an enzyme fundamental in miRNA biogenesis, have shown that $Dicer^{-/-}$ females are sterile and display multiple reproductive defects [22]. In pigs, the miRNA transcriptome has been profiled in adult ovary where a total of 451 miRNAs have been described and 224 unique miRNAs were differential expressed between ovary and testis, most of them located on porcine chromosome X (X-linked miRNAs) [23]. Moreover, in porcine placenta, differential miRNA expression between day 30 and 90 of gestation has been determined reflecting the importance of miRNAs during porcine placenta development [24]. Finally, a polymorphism has been identified in miR-27a which has been associated with litter size [25].

In the present work, we have profiled miRNA expression in the ovaries of pregnants sows with different performances regarding the number of embryos attached to the uterus at day 30-32 of gestation. Sows from this study belong to an experimental F_2

intercross between *Iberian* (Ib) and *Meishan* (Me) breeds for which data on QTL for litter size [26] and differentia gene expression from the ovary of pregnant sows [27] are available. The two parental breeds (Ib and Me) are highly divergent at the genetic and phenotypic levels [28, 29] displaying extreme phenotypes for prolificacy traits making them suitable for performing prolificacy studies. Meishan is considered one of the most prolific breed whereas Iberian is one very low prolificacy breed. This superiority is due to higher ovulation rates and reduced embryo mortality [30]. These results suggest that differences in prolificacy, determined by the litter size, might also be regulated by genes expressed in the ovary during the gestational period.

RESULTS

miRNA expression analysis

miRNA expression profile was evaluated through miRCURY LNA™ microRNA microarrays v.11.0 (Exiqon, Denmark) in ovary of pregnant sows displaying extreme phenotypes for the number of embryos (NE) attached to the uterus. According the NE, sows were divided into two groups: high prolificacy sows (NE≥ 16; n = 3) and low prolificacy sows (NE≤ 6; n = 3). All hybridised chips passed the quality control and were, therefore, included in the analysis. After data normalisation, filtration by low variance (Var > 0.02) and exclusion of probes with an absolute Log_2Difference ≤ 0.2 between high and low prolificacy groups, 84 probes were left for further analysis (Additional file 1). None of the probes passed the statistical significance threshold of FDR < 0.05 after the t-test analysis. However, these probes were able to completely separate the high from the low prolificacy sows by principal component analysis (Figure 1). Thirty-four out of the 84 probes could not be annotated in the pig genome as miRNAs because they were not considered miRNAs or were not homologous to any region of the pig genome (*Sus scrofa*, Ensembl release 64 - September 2011). The remaining 50 probes were annotated as miRNAs (Table 1), representing 41 unique miRNAs. For eight miRNAs, two or three different replicate probes representing the same miRNA appeared on the microarray. All the hybridisation signals were in agreement in the direction of the miRNA gene expression, showing the robustness of

the microarray hybridisation. Twenty-nine miRNAs had already been described in pigs and 12 were orthologous miRNAs and were, therefore, considered as new pig miRNAs (miRBase v.18). From the 41 miRNAs, thirty-six miRNAs were more expressed in high prolificacy sows whereas five were less expressed (Table 1).

The microarray data revealed that expression differences between high and low prolificacy sows are low (Log_2Differences ≤ 0.58). These results were not surprising due to the low sample size (n = 6) and the fact that sows have similar genetic background because they come from the same F_2 intercross although they belong to different litters. However, phenotypic differences between the two prolificacy groups were relatively high (mean difference of twelve embryos between high and low prolificacy groups). Taking this into account, we selected as differentially expressed (DE) the 5 miRNAs showing maximal expression differences regarding the prolificacy levels and with a nominal $P < 0.20$. The four highest upregulated miRNAs in high prolificacy sows (miR-146a-5p, miR-142-3p, miR-142-5p, miR-146b) (Log_2Difference ≥ 0.51; $P \leq 0.12$; Table 1) and the highest downregulated miRNA (miR-335) in high prolificacy sows (Log_2Differences = 0.58; $P < 0.20$; Table 1) were selected for validation by RT-qPCR.

RT-qPCR validation of the differentially expressed miRNAs

The five miRNAs displaying maximal differences between high and low prolificacy sows were selected for validation by RT-qPCR in order to measure the miRNA expression in a more accurate way. Reference miRNAs (miR-103, miR-17-3p, and let-7a) were chosen based on their expression stability (M < 1.5) according to the microarray expression data from the present study. Porcine specific sequences (miRBase v.18) were used for primer design in order to reduce the error in miRNA expression quantification due to the fact that RT-qPCR is highly sequence specific. Initially, RT-qPCR was performed in the same samples used for microarray experiments (n=6). Good agreement between the array and the qRT-PCR data was obtained (Pearson correlation coefficient (R^2) ≥ 0.77) except for miR-142-5p that was very low (R^2 = 0.17; Table 2). Afterwards, to confirm the results of RT-qPCR validation of DE miRNAs, the sample size was expanded (n = 35; 15 from high prolificacy and 20 from low prolificacy). miR-146a-5p was still

significantly differentially expressed (*P* = 0.02, 1.5-fold-increased) and miR-142-3p was very close to be differentially expressed (*P* = 0.07, 1.2 fold-increased) in high *versus* low prolificacy sows (Table 2). Both miRNAs were denominated prolificacy-associated miRNAs.

mRNA target prediction and functional analyses for prolificacy miRNAs

mRNA targets for the prolificacy-associated miRNAs (miR-146a-5p and miR-142-3p) were predicted *in silico* using the Ingenuity Pathways Analysis software (Ingenuity® Systems, USA). Only experimentally verified and highly predicted gene targets were considered for the analysis. A total of 628 mRNAs were identified as targets for the two differentially expressed miRNAs (Additional file 2). Ten mRNAs were common targets for both prolificacy miRNAs whereas 325 and 293 were specific mRNA targets for miR-146a-5p and miR-142-3p, respectively (Additional file 2).

Functional annotation clustering of mRNA targets with DAVID 6.7 software [31, 32] helped in determining the biological role of the prolificacy-associated miRNAs (Table 3). Nine GO clusters were found for the miR-146a-5p target mRNAs, six of them were related to immune system processes, such as responses to stimulus or bacterium, regulation of immune response and cytokine production and chemotaxis, and the remaining GO clusters were related to metabolic and developmental processes (Table 3). For miR-142-3p, five clusters were determined which were associated with the regulation of transcription, intracellular and protein transport, cell communication, regulation of cellular and metabolic processes and cell cycle (Table 4).

As previously mentioned, the ovarian mRNA transcriptional profile of high and low prolificacy sows used in this work was established in a previous study [27]. The Porcine Genome Array (Affymetrix, Santa Clara, CA, USA) was used and a total of 290 differentially expressed genes (FDR < 0.1) were identified. Among these 290 differentially expressed genes, 11 genes could be targets for prolificacy-related miRNAs found in the current study (Table 5). Eight differentially expressed genes were predicted to be targeted by miR-146a-5p (*BTG2*, *CCL8*, *LRRK1*, *TMSB15A*, *CFH*, *FAF2*,

PPM1K and *TLR1*) and three by miR-142-3p (*CPEB2*, *SLC35F5* and *DDX39*). miRNAs regulate the expression level of their mRNA target through mRNA degradation or by repression of the mRNA translation. In this sense, the correlations between prolificacy miRNAs and their DE target mRNAs were calculated (Table 5). High negative correlations were determined between miR-146a-5p and *PPM1K* (P = 0.003) and *TLR1* (P = 0.17) and between miR-142-3p and *CPEB2* (P = 0.032) (Table 5). Thus, in high prolificacy sows, prolificacy-associated miRNAs were upregulated whereas *PPM1K, TLR1*and *CPEB2* were downregulated.

DISCUSSION

In this study, we have compared the miRNA expression profiles of ovaries from pregnant IbxMe F_2 sows displaying extreme phenotypes regarding the prolificacy levels (defined as the number of embryos attached to the uterus at 30-32 days of gestation). Total RNA from ovary samples from three high prolificacy sows (NE ≥ 16) and three low prolificacy sows (NE ≤ 6) were profiled for miRNA gene expression using t he miRCURY LNA™ microRNA Microarray (Exiqon, Denmark). Due to the low variability in miRNA gene expression found between high and low prolificacy groups, the usual stringent significance thresholds applied to microarray studies could not be implemented to our data. To assess differential miRNA expression between the high and low prolificacy groups, unsupervised classification methodologies were used followed by a conventional t-test analysis to assess the fold change between groups. Differentially expressed miRNAs with most differential expression between high and low prolificacy sows were selected although the significance threshold was relatively high (P < 0.2). It was due to the fact that microarray data revealed very low differences in miRNA expression between the two groups. This result might not be surprising giving the fact that one could not expect to find large expression differences only depending on the reproductive performance of the sows, but slight differences in the expression level of a regulatory factor could lead to big differences at the phenotypic level. Furthermore, the microarrays technique is not accurate enough to assess for small differences in

gene expression [33] and the sample size might be higher in order acquire to sufficient power to detect small expression differences.

Overall, we found 36 miRNAs to be upregulated in high prolificacy sows, whereas only 5 miRNAs were upregulated in the low prolificacy group. Most of the miRNAs detected in microarrays had been reported previously to be expressed in ovary of pigs and/or mice [23]. However, this is the first time that miR-223-3p, miR-711-3p, miR-93-5p and miR-18b are described in the ovary of sows. Moreover, 8 miRNAs (miR-106a, miR-17, miR-20, miR-24, miR-125b, miR-92b, miR-27a and let-7i) were shown to be differentially expressed in porcine placenta between days 30 and 90 of gestation suggesting a potential role of these miRNAs in porcine placental growth and function [24]. Interestingly, miRNAs upregulated in placenta at 30 days (miR-106a, miR-17 and miR-20) and miRNAs upregulated in placenta at 90 days (miR-27a, miR-125b, miR-92b and let-7i) were expressed in the ovary of pregnant sows being all upregulated in high prolificacy sows unless miR-125b was downregulated. However, the ovarian differential expression between high and low prolificacy sows was smaller than the placental expression at day 30 and 90 of gestation. miR-27a was differentially expressed from Large White ovaries to Chinese Erhualian ovaries and a SNP mutation in the miR-27a, located at 18bp downstream of pre-miR-27a, has been associated to a higher litter size in DIV pigs population (4^{th} Dam line of Chinese lean-type new lines) [25]. Nonetheless, our study revealed that the difference in miR-27a expression regarding the prolificacy levels of pregnant Ib x Me F_2 sows was small. Further analyses are needed in order to determine if the SNP described in miR-27a is segregating in our population and if it could explain the difference in miR-27a expression levels between high and low prolificacy sows.

Five miRNAs displaying maximal differences in gene expression were validated by RT-qPCR using the same samples as in the microarray analysis. Importantly, all five miRNAs are located on porcine chromosomes where epistatic QTL interaction for prolificacy traits in the same experimental samples have been described [26]. miR-146a-5p, miR-142-3p and miR-142-5p, miR-146b and miR-335 are located on porcine chromosome 16 (SSC16), SSC12, SSC14 and SSC18, respectively (*Sus scrofa*, Ensembl release 66 - February 2012). All except one (miR-142-5p) showed a high level of

concordance of expression level between microarray and RT-qPCR data ($R \geq 0.77$). The lack of correlation found for miR-142-5p ($R = 0.17$) could be attributed to the higher noise generated by using an array including non-specific probes or by cross hybridisation [33]. To verify that these miRNAs were really differentially expressed due to the prolificacy level, we expanded the sample size to n = 35, including 15 sows from high prolificacy and 20 from low prolificacy. Overall, two miRNAs (miR-146a-5p and miR-142-3p) remained significant and were, therefore, called prolificacy-associated miRNAs. Differential expression was observed between miR-146a-5p and miR-146b even though similar expression pattern is expected as they belong to the same miRNA family. It is still unclear if these two miRNAs have redundant and/or different functions [34]. Nonetheless, small fold changes in miRNAs expression may be biologically significant due to the fact that miRNAs regulate many different mRNA targets which could amplify their effect and may result in large phenotypic differences [35]. Both prolificacy-associated miRNAs have been shown to play key roles in reproductive processes in humans and rodents. miR-146a-5p has been associated with oocyte maturation, embryo development at preimplantation time [36, 37] and with fetal growth in placenta [37]; whereas miR-142-3p is expressed in human chorioamniotic membranes at 37 weeks of gestation [38] and has been associated with endometriosis [21]. miR-146a-5p is located on porcine chromosome 16 (SSC16) and miR-142-3p on SSC12 (*Sus scrofa*, Ensembl release 66 - February 2012). Interestingly, epistatic QTL have been described in both chromosomes [26] and, moreover, on SSC12, two prolificacy QTL have been reported [39]. As miRNAs are regulatory molecules, they could regulate genes in *cis* which could be candidates for the prolificacy QTLs described by [26, 39] . The identification of the targets for prolificacy-associated miRNAs located on prolificacy QTL would be interesting in order to detect new candidate genes that could be used for the improvement of reproductive traits in pigs.

The identification of the targets and their functional importance is the greatest challenge to understand the biological role of miRNA expression. A total of 628 target mRNAs were predicted for the two prolificacy-associated miRNAs. *In silico* functional analyses of miR-146a-5p mRNA targets suggest that the targeted genes are mainly involved in regulating the expression of genes related to the immune system

processes. In particular, miR-146a-5p could act as a tuning mechanism that maintains the immune homeostasis by limiting the inflammatory response [34]. The immune system response during the gestation process is very important for the establishment of maternal-foetal tolerance and it is also critical for the embryo implantation and maintenance of pregnancy [40]. During pregnancy, the immune system is self-regulated in reproductive tissues to avoid the rejection of the foetus and tolerate foetal alloantigens [41, 42]. Specifically, it has been postulated by Pate *et al.,* (2001) [43] that immune cells in the ovary have a role in abating the inflammatory response against foetal antigens. Our results suggest that higher levels of miR-146a-5p expression could lead to a better efficiency of immunosuppression which is essential for materno-foetal tolerance.

miR-142-3p has been involved in the induction of matrix metalloproteases (MMP). They are important in the cell-to-extracellular matrix (ECM) interactions which are essential for cell survival and migration [44]. Functional annotation of miR-142-3p mRNA targets also suggests an important role of this miRNA in the regulation of transcription, transport and cell communication processes and in the regulation of metabolic processes. Importantly, these processes would contribute to homeostasis maintenance, necessary for a correct functional development of the ovarian tissue [45] which has extremely high metabolic activity during pregnancy.

The comparison of the mRNA targets of the prolificacy-associated DE miRNAs with the DE mRNAs obtained by [27] in the same population has allowed us to identify a number of DE mRNAs that could be directly regulated by the prolificacy-associated miRNAs. The expression levels of miR-146a-5p and miR-142-3p were negatively correlated with the expression of the DE mRNAs *PPM1K* and *TLR1,* and *CPEB2*, respectively. However, we could not discard the effect of prolificacy-related miRNAs on the other genes due to the lack in the correlation detection could be associated to the fact that miRNAs act as fine-tunning mechanisms and the increase/decrease level of mRNA expression could not be detected at time of measure. *PPM1K* is a mitochondrial protein phosphatase which is essential for cell survival and development [46]. Although loss of *PPM1K* expression is associated with cell death [46], results showed that downregulation of *PPM1K* would benefit in increasing the prolificacy level

in sows [27]. *TLR1* is a class of pattern recognition receptor (PRR) which participates in innate immunity and is expressed in various gestation-associated tissues. Activation of TLRs results in an inflammatory immune response [47] that facilitates the adaptive immune response [48]. It is considered that an excessive or aberrant inflammatory response is generally harmful to pregnancy [49]. Moreover, it is hypothesised that sows with lower immune system activation are less prone to materno-foetal rejection [27]. *CPEB2* is a key regulator of local protein synthesis by regulating mRNA translation through changes in the polyA tail length [45, 50]. Thus, it seems that it participates in cellular maintenance which is important for ovarian function [45]. This result is in agreement with the main role associated with miR-142-3p determined by functional annotation. Overall, we could hypothesise that *PPM1K*, *TLR* and *CPEB2* could be regulated by a novel post-transcriptional regulation mechanism mediated by the prolificacy-associated miRNAs, miR-146a-5p and miR-142-3p, in IbxMe F_2 sows. However, further studies must be performed in order to confirm these results.

CONCLUSIONS

miR-146a-5p and miR-142-3p have been associated, for the first time, with prolificacy levels in IbxMe F_2 sows. Our results suggest that both prolificacy-associated miRNAs are important for reproduction processes. Specifically, they are involved in the immune system and in cellular maintenance providing a uterine environment favorable to embryonic development.

METHODS

Animal material

Animals used in this study come from an IbxMe F_2 intercross described thoroughly in [26]. Briefly, 3 Ib boars (*Guadyerbas* line) were mated with 18 Me sows to obtain 255 F_2 reproductive sows. Four consecutive parities were recorded and in the 5^{th} gestation sows were slaughtered at 30-32 days of pregnancy. The number of embryos (NE) was recorded and samples from ovarian tissue were collected, snap frozen in liquid

nitrogen and stored at -80ºC until usage. Sows displaying extreme phenotypes regarding the NE were selected and classified as high or low prolificacy.

Total RNA extraction

Total RNA was extracted using the TRIzol® reagent (Invitrogen) following the manufacturer's instructions. RNA quality was evaluated with the RNA 6000 Nano Kits on an Agilent 2100 Bioanalyzer (Agilent, Germany) and the RNA quantity was assessed using Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA).

Microarray hybridisation and scanning

For microarrays studies total RNA from ovary samples from 3 high (NE≥16) and 3 low (NE≤6) IbxMe F_2 sows were used. One microgram of total RNA from each sample was labeled using the miRCURY™ Power Labeling Kit (Exiqon, Denmark) following manufacturer's specifications. Labeled RNA was hybridised for 16 hours at 65 °C to miRCURY LNA™ microRNA microarrays v.11.0 (Exiqon, Denmark) which contain 1891 capture probes in quadruplicate, targeting miRNAs from human, mouse and rat registered in the miRBase v14.0. After washing and drying, the microarrays were scanned in a DNA Microarray Scanner (Agilent) and the resulting images were quantified using Imagene v. 8.0 (BioDiscovery, CA).

Microarray data analysis

Inter-slide normalisation was performed using a non-linear quantile-based normalisation algorithm in Rosetta Resolver (Rosetta Biosoftware). Following normalisation, data was Log_2 transformed. Two filtering steps were conducted. First, a variance filter was applied to remove all signals with variance below 0.02, and then, filtering for an absolute Log_2Difference > 0.2 (corresponding to a 15% difference in signal) between low and high prolificacy groups was applied. Overall, 84 probe signals were left for further analysis. Student's t-test was applied to test for differentially

expressed miRNAs between low and high prolificacy groups. Heat maps and two-way hierarchical clustering were based on Pearson correlation coefficients and performed in Qlucore Omics Explorer v.2.2 (Qlucore, Lund, Sweden). The 84 probes with more variable expression between the two groups were manually checked in miRBase version 18 and then annotated to the pig genome (*Sus scrofa*, Ensembl release 64 - September 2011).

RT-qPCR

Five miRNAs (miR-146a-5p, miR- 142-3p, miR-142-5p, miR146b and miR-335) were selected as candidate miRNAs to be evaluated through RT-qPCR. Reference miRNAs (miR-103, miR-17-3p and let-7a) were chosen based on their stability according to microarray data obtained in the present study.

For RT-qPCR analyses we used total RNA from ovary samples from 15 high (N≥ 13) and 20 low (N≤ 11) prolificacy sows (including the 6 samples employed in the microarray analysis). Reverse transcription reaction was performed as described by [51] from 400 ng of total RNA. Two technical replicates of each RNA sample were used for reverse transcription.

Quantitative PCR reactions were performed in duplicate for each synthesised cDNA in 20 µl total volume with 5µl of cDNA diluted 20 times, 10 µl 1x Power SYBR green PCR Master Mix (Applied Biosystems) and 0.25µM of each primer (Table 6) on an ABI PRISM® 7900 sequence detection system (Applied Biosystems, Warrington, UK).

Primers used in qPCR were synthesised following the rules described by [51]. qPCR efficiency for each miRNA assay was calculated using standard curves with 10-fold dilutions (made with equal amount of cDNA from all the samples). Thermal profile was 95 ºC for 10 min and 40 cycles of 95 ºC for 15 sec and 60 ºC for 60 sec. Melting curve analysis was performed after each thermal qPCR profile to ensure the specificity of the assay.

RT-qPCR Data Analysis

GeNorm algorithm [52] was used to evaluate the stability of the reference miRNAs (miR-103, miR-17-3p and let-7a) used (M<1.5) and to obtain a normalisation factor (NF) based on the combined information of all reference miRNAs used. Data from RT-qPCR was analysed using quantities obtained from quantification cycle (Cq) and using standard curve in order to consider amplification efficiencies. Quantities for each sample were normalised by NF and fold changes were then calculated based on the lowest normalised quantity. Fold changes were Log_2 transformed and used for statistical analysis. Differential miRNA expression was performed using an analysis of variances (ANOVA) taking two factors into account: retrotranscription's replicate (RT-1 and RT-2) and prolificacy level (High and Low). Data analysis was performed using the PROC GLM of the Statistical Analysis System (Statistics, V 9.1.3; SAS Institute, Inc., Cary, NC). Significance threshold was set at $\alpha < 0.05$.

Gene Target Identification and Functional Analysis

Putative target mRNAs for differentially expressed miRNAs (miR-146a-5p and miR-142-3p) were identified with the *Ingenuity Pathway Analysis* (IPA) software (Ingenuity® Systems, USA). Only targets that were experimentally observed and highly predicted were selected in order to reduce the number of false positive hits. Selected miRNAs targets were functionally annotated through Gene Ontology (GO) analysis with DAVIDGO 6.7 [31, 32] by using the functional annotation clustering option that allow clustering the annotation terms based on the genes included in each annotation term. Analyses were performed by using the human annotation with DAVID software (http://david.abcc.ncifcrf.gov/home.jsp) because the human genome is better annotated than the porcine genome leading to a more accurate functional analysis. Correction for multiple testing was done by using the Benjamini–Hochberg method [52]. Significant threshold was set at $\alpha < 0.05$.

AUTHORS' CONTRIBUTIONS

IB, SC, AS and AT designed the study. JLN and AS participated in the design of Iberian x Meishan F_2 intercross. AF, JLN, AS and AT participated in the experimental protocol and samples recollection. IB and OT prepared RNA samples. TL and RS performed microarray experiments and analysed them. IB carried out RT-qPCR experiments, mRNA target identification and functional analysis, analysed the results and interpreted the data. All authors participated in writing the manuscript and revised it critically from a draft by IB.

ACKNOWLEDGEMENTS

The authors are indebted to Luis Varona, M. Arqué, J. Tarrés, M. Fina, and the staff of Nova Genètica, in particular to E. Ramells, F. Márquez, R. Malé, F. Rovira, and I. Riart, for cooperating in the experimental protocol. The authors gratefully acknowledge INRA (France) and the CIA El Dehesón del Encinar (Spain) for providing the purebred Meishan sows and Iberian boars, respectively. The authors thank Peter K Busk for RT-qPCR primer design and Sara Collins for helping in performing the RT-qPCR experiments. This research was funded in part by Project AGL2004-08368-C03, AGL2007-66371-C02-01, AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 Program (CSD2007-00036), both from the Spanish Ministry of Science and Innovation. IB is recipient of PIF PhD fellowship from Universitat Autònoma de Barcelona.

REFERENCES

1. Pope, WF: Embryonic mortality in swine. In Embryonic Mortality in Domestic Species. 1st edition. Edited by CRC Press. Florida: Boca Raton; 1994:53–77

2. Spotter A, Distl O: Genetic approaches to the improvement of fertility traits in the pig. Vet J 2006, 172:234-247

3. Distl O: Mechanisms of regulation of litter size in pigs on the genome level. Reprod Domest Anim 2007, 42:10-16

4. Brevini TAL, Cillo F, Antonini S, Tosetti V, Gandolfi F: Temporal and spatial control of gene expression in early embryos of farm animals. Reprod Fertil Dev 2007, 19:35-42

5. Ostrup E, Bauersachs S, Blum H, Wolf E, Hyttel P: Differential endometrial gene expression in pregnant and nonpregnant sows. Biol Reprod 2010, 83:277-285

6. Ka H, Seo H, Kim M, Moon S, Kim H, Lee CK: Gene expression profiling of the uterus with embryos cloned by somatic cell nuclear transfer on day 30 of pregnancy. Anim Reprod Sci 2008, 108:79-91

7. Ross JW, Ashworth MD, White FJ, Johnson GA, Ayoubi PJ, DeSilva U, Whitworth KM, Prather RS, Geisert RD: Premature estrogen exposure alters endometrial gene expression to disrupt pregnancy in the pig. Endocrinology 2007, 148:4761-4773

8. Hunter MG, Robinson RS, Mann GE, Webb R: Endocrine and paracrine control of follicular development and ovulation rate in farm species. Anim Reprod Sci 2004, 82-83:461-477

9. Waclawik A: Novel insights into the mechanisms of pregnancy establishment: regulation of prostaglandin synthesis and signaling in the pig. Reproduction 2011, 142:389-99

10. Bell JT, Spector TD: A twin approach to unraveling epigenetics. Trends Genet 2011, 27:116-125

11. Hawkins SM, Buchold GM, Matzuk MM: Minireview: The Roles of Small RNA Pathways in Reproductive Medicine. Mol Endocrinol 2011, 25:1257-1279

12. Morales Prieto DM, Markert UR: MicroRNAs in pregnancy. J Reprod Immunol 2011, 88:106-111

13. Carletti MZ, Christenson LK: MicroRNA in the ovary and female reproductive tract. J Anim Sci 2009, 87:E29-38

14. Baley J, Li J: MicroRNAs and ovarian function. J Ovarian Res 2012, 5:8

15. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ: A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature 2010, 465:584-589

16. Pan Q, Chegini N: MicroRNA signature and regulatory functions in the endometrium during normal and disease states. Semin Reprod Med 2008, 26:479-493

17. Chakrabarty A, Tranguch S, Daikoku T, Jensen K, Furneaux H, Dey SK: MicroRNA regulation of cyclooxygenase-2 during embryo implantation. Proc Natl Acad Sci U S A 2007, 104:15144-15149

18. Cochrane DR, Cittelly DM, Richer JK: Steroid receptors and microRNAs: relationships revealed. Steroids 2011, 76:1-10

19. Zhou Y, Zhu YZ, Zhang SH, Wang HM, Wang SY, Yang XK: MicroRNA expression profiles in premature ovarian failure patients and its potential regulate functions. Chinese Journal of birth health and heredity 2011, 19:20-22

20. Mouillet JF, Chu T, Sadovsky Y: Expression patterns of placental microRNAs. Birth Defects Res A Clin Mol Teratol 2011, 91:737-743

21. Ohlsson Teague EM, Van der Hoek KH, Van der Hoek MB, Perry N, Wagaarachchi P, Robertson SA, Print CG, Hull LM: MicroRNA-regulated pathways associated with endometriosis. Mol Endocrinol 2009, 23:265-275

22. Nagaraja AK, Andreu-Vieyra C, Franco HL, Ma L, Chen R, Han DY, Zhu H, Agno JE, Gunaratne PH, DeMayo FJ, Matzuk MM: Deletion of Dicer in somatic cells of the female reproductive tract causes sterility. Mol Endocrinol 2008, 22:2336-2352

23.Li M, Liu Y, Wang T, Guan J, Luo Z, Chen H, Wang X, Chen L, Ma J, Mu Z, Jiang AA, Zhu L, Lang Q, Zhou X, Wang J, Zeng W, Li N, Li K, Gao X, Li X: Repertoire of porcine microRNAs in adult ovary and testis by deep sequencing. Int J Biol Sci 2011, 7:1045-1055

24. Su L, Zhao S, Zhu M, Yu M: Differential expression of microRNAs in porcine placentas on Days 30 and 90 of gestation. Reprod Fertil Dev 2010, 22:1175-1182

25. Lei B, Gao S, Luo LF, Xia XY, Jiang SW, Deng CY, Xiong YZ, Li FE: A SNP in the miR-27a gene is associated with litter size in pigs. Mol Biol Rep 2011, 38:3725-3729

26. Noguera JL, Rodriguez C, Varona L, Tomas A, Munoz G, Ramirez O, Barragan C, Arque M, Bidanel JP, Amills M, Ovilo C, Sanchez A: A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. BMC Genomics 2009, 10:636

27. Fernandez-Rodriguez A, Munoz M, Fernandez A, Pena RN, Tomas A, Noguera JL, Ovilo C, Fernandez AI: Differential gene expression in ovaries of pregnant pigs with high and low prolificacy levels and identification of candidate genes for litter size. Biol Reprod 2011, 84:299-307

28. Ramirez O, Ojeda A, Tomas A, Gallardo D, Huang LS, Folch JM, Clop A, Sanchez A, Badaoui B, Hanotte O, Galman-Omitogun O, Makuza SM, Soto H, Cadillo J, Kelly L, Cho IC, Yeghoyan S, Perez-Enciso M, Amills M: Integrating Y-chromosome,

mitochondrial, and autosomal data to analyze the origin of pig breeds. Mol Biol Evol 2009, 26:2061-2072

29. Alves E, Ovilo C, Rodriguez MC, Silio L: Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. Anim Genet 2003, 34:319-324

30. Ashworth, CJ: Advances in embryo mortality research. In Proceedings of the International Pig Veterinary Society, 1998, 1:231–237

31. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009, 4:44-57

32. Huang DW, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 2009, 37:1-13

33. Chuaqui RF, Bonner RF, Best CJ, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M, Linehan WM, Knezevic V, Emmert-Buck MR: Post-analysis follow-up and validation of microarray experiments. Nat Genet 2002, 32:509-514

34. Rusca N, Monticelli S: MiR-146a in Immunity and Disease. Mol Biol Int 2011, 2011:437301

35. Calin GA, Croce CM: MicroRNA-cancer connection: the beginning of a new tale. Cancer Res 2006, 66:7390-7394

36. Abd El Naby WS, Hagos TH, Hossain MM, Salilew-Wondim D, Gad AY, Rings F, Cinar MU, Tholen E, Looft C, Schellander K, Hoelker M, Tesfaye D: Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. Zygote 2011, 11:1-21

37. Maccani MA, Padbury JF, Marsit CJ: miR-16 and miR-21 expression in the placenta is associated with fetal growth. PLoS One 2011, 6:e21210

38. Montenegro D, Romero R, Kim SS, Tarca AL, Draghici S, Kusanovic JP, Kim JS, Lee DC, Erez O, Gotsch F, Hassan SS, Kim CJ: Expression patterns of microRNAs in the chorioamniotic membranes: a role for microRNAs in human pregnancy and parturition. J Pathol 2009, 217:113-121

39. Fernandez-Rodriguez A, Rodriguez C, Varona L, Balcells I, Noguera JL, Ovilo C, Fernandez AI: Analysis of candidate genes underlying two epistatic quantitative trait loci on SSC12 affecting litter size in pig. Anim Genet 2010, 41:73-80

40. Mor G, Abrahams VM: Potential role of macrophages as immunoregulators of pregnancy. Reprod Biol Endocrinol 2003, 1:119

41. Pate JL, Toyokawa K, Walusimbi S, Brzezicka E: The interface of the immune and reproductive systems in the ovary: lessons learned from the corpus luteum of domestic animal models. Am J Reprod Immunol 2010, 64:275-286

42. Guerin LR, Prins JR, Robertson SA: Regulatory T-cells and immune tolerance in pregnancy: a new target for infertility treatment? Hum Reprod Update 2009, 15:517-535

43. Pate JL, Landis Keyes P: Immune cells in the corpus luteum: friends or foes? Reproduction 2001, 122:665-676

44. Kim D, Song J, Kim S, Kang SS, Jin EJ: MicroRNA-142-3p regulates TGF-beta3-mediated region-dependent chondrogenesis by regulating ADAM9. Biochem Biophys Res Commun 2011, 414:653-659

45. Bukovsky A, Caudle MR, Keenan JA: Dominant role of monocytes in control of tissue function and aging. Med Hypotheses 2000, 55:337-347

46. Lu G, Ren S, Korge P, Choi J, Dong Y, Weiss J, Koehler C, Chen JN, Wang Y: A novel mitochondrial matrix serine/threonine protein phosphatase regulates the mitochondria permeability transition pore and is essential for cellular survival and development. Genes Dev 2007, 21:784-796

47. Medzhitov R, Janeway CA,Jr: Innate immunity: impact on the adaptive immune response. Curr Opin Immunol 1997, 9:4-9

48. Medzhitov R, Preston-Hurlburt P, Janeway CA,Jr: A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. Nature 1997, 388:394-397

49. Patni S, Flynn P, Wynen LP, Seager AL, Morgan G, White JO, Thornton CA: An introduction to Toll-like receptors and their possible role in the initiation of labour. BJOG 2007, 114:1326-1334

50. Kurihara Y, Tokuriki M, Myojin R, Hori T, Kuroiwa A, Matsuda Y, Sakurai T, Kimura M, Hecht NB, Uesugi S: CPEB2, a novel putative translational regulator in mouse haploid germ cells. Biol Reprod 2003, 69:261-268

51. Balcells I, Cirera S, Busk PK: Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. BMC Biotechnol 2011, 11:70

52. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: Accurate normalization of real-time quantitative RT-PCR data by

geometric averaging of multiple internal control genes. Genome Biol 2002, 3:RESEARCH0034.

53. Benjamini Y, Hochberg Y: Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 1995, 57:289-300

TABLES

Table 1. Microarray expression profile of miRNAs in ovary of IbxMe F_2 pregnant sows regarding prolificacy levels.

Porcine miRNA Annotation	Log_2 Mean Expression High Prolificacy	Log_2Mean Expression Low Prolificacy	Log_2Mean Difference (High-Low)	Log_2Fold Change High/Low Prolificacy[a]	Mean Variance	*P-value*[b]
ssc-miR-146a-5p	**8.19**	**7.53**	**0.66**	**1.09**	**0.182**	**0.078**
ssc-miR-142-3p	**10.05**	**9.48**	**0.58**	**1.06**	**0.122**	**0.013**
ssc-miR-142-5p	**8.84**	**8.32**	**0.53**	**1.06**	**0.127**	**0.054**
ssc-miR-146b	**8.50**	**7.99**	**0.51**	**1.06**	**0.132**	**0.118**
ssc-miR-452-5p	8.47	8.04	0.43	1.05	0.122	0.194
ssc-miR-223-3p	8.69	8.27	0.42	1.05	0.090	0.098
ssc-miR-144-5p	6.90	6.49	0.41	1.06	0.190	0.333
ssc-miR-106a	10.03	9.68	0.35	1.04	0.058	0.074
ssc-miR-20a	9.73	9.39	0.34	1.04	0.063	0.117
ssc-miR-452-5p	7.32	6.98	0.34	1.05	0.068	0.178
ssc-miR-16	12.39	12.06	0.33	1.03	0.042	0.041
ssc-miR-221	8.53	8.20	0.33	1.04	0.050	0.067
ssc-miR-222	9.11	8.79	0.33	1.04	0.067	0.154
ssc-miR-93-5p	9.32	9.00	0.32	1.04	0.035	0.009
ssc-miR-15a	10.83	10.52	0.31	1.03	0.042	0.087
ssc-miR-20b-5p	8.62	8.31	0.30	1.04	0.046	0.073
ssc-miR-132-3p	8.44	8.15	0.30	1.04	0.045	0.130
ssc-miR-18b	7.22	6.93	0.29	1.04	0.033	0.023
ssc-miR-129a	6.61	6.32	0.29	1.05	0.031	0.021
ssc-miR-212-3p	8.44	8.15	0.28	1.03	0.042	0.132
ssc-mir-27a	10.34	10.07	0.28	1.03	0.037	0.078
ssc-let-7f	10.57	10.30	0.28	1.03	0.034	0.056
ssc-miR-17	10.03	9.78	0.26	1.03	0.045	0.152
ssc-miR-129b-3p	10.03	9.78	0.26	1.03	0.067	0.267
ssc-miR-30a-5p	9.54	9.28	0.26	1.03	0.029	0.048
ssc-miR-342	7.38	7.13	0.25	1.03	0.033	0.086
ssc-miR-18b	7.99	7.75	0.24	1.03	0.040	0.204
ssc-miR-92a	9.20	8.97	0.24	1.03	0.038	0.204
ssc-miR-21	13.67	13.44	0.23	1.02	0.026	0.108
ssc-miR-132-5p	8.05	7.82	0.23	1.03	0.020	0.060
ssc-let-7i	11.29	11.06	0.23	1.02	0.023	0.074
ssc-miR-665-3p	7.30	7.07	0.23	1.03	0.236	0.627
ssc-miR-20b-5p	8.01	7.79	0.22	1.03	0.020	0.075
ssc-miR-199a-3p	11.05	10.82	0.22	1.02	0.162	0.581
ssc-miR-106a	9.49	9.27	0.22	1.02	0.025	0.087
ssc-miR-582-3p	11.04	10.82	0.22	1.02	0.029	0.156
ssc-miR-15b	10.20	9.98	0.22	1.02	0.024	0.075
ssc-miR-21	10.00	9.78	0.22	1.02	0.032	0.151
ssc-miR-18b	8.12	7.91	0.21	1.03	0.038	0.269
ssc-miR-28-5p	7.85	7.64	0.21	1.03	0.041	0.257
ssc-miR-92b-5p	7.28	7.06	0.21	1.03	0.062	0.371
ssc-miR-17	9.79	9.59	0.21	1.02	0.027	0.140
ssc-miR-143	13.86	13.65	0.21	1.02	0.058	0.361
ssc-miR-155	7.72	7.92	-0.20	0.98	0.221	0.663
ssc-miR-125b	13.00	13.20	-0.20	0.98	0.176	0.639
ssc-miR-711-3p	6.96	7.23	-0.27	0.96	0.030	0.031

ssc-miR-135-5p	7.86	8.15	-0.29	0.96	0.082	0.263
ssc-miR-711-3p	6.54	6.87	-0.33	0.95	0.048	0.079
ssc-miR-135-5p	8.00	8.38	-0.38	0.95	0.111	0.200
ssc-miR-335	**10.47**	**11.05**	**-0.58**	**0.95**	**0.250**	**0.181**

In bold, the miRNAs displaying maximal differences between high and low prolificacy sows.
[a] Positive and negative signs indicate that the level of gene expression is higher for the first or the second group of the test, respectively.
[b] Nominal *P-value*. Significance was set at 0.05.

Table 2. RT-qPCR expression of selected miRNAs in pregnant ovary of IbxMe F2 sows regarding prolificacy levels.

miRNA	Microarray Validation (n = 6)			Further validation (n = 35)	
	Fold Change High / Low Prolificacy[a]	*P-Value*[b]	R^{2c}	Fold Change High / Low Prolificacy[a]	*P-Value*[b]
miR-146a-5p	1.09	**0.08**	0.87	1.47	**0.02**
miR-142-3p	1.06	**0.01**	0.97	1.17	**0.07**
miR-146b	1.06	0.12	0.77	1.15	0.42
miR-335	-1.06	0.18	0.93	-1.05	0.72
miR-142-5p	1.06	**0.05**	0.17	1.00	0.89

[a] Positive and negative signs indicate that the level of gene expression is higher for the first or the second group of the test, respectively.
[b] Nominal *P-values* testing the prolificacy effect. Significance was set at $P < 0.05$
[c] Correlation (Pearson coefficient) of the microarray expression and RT-qPCR expression considering the same animal size (n = 6) as used for microarray experiment.

Table 3. Functional annotation cluster for candidate mRNAs targets for miR-146a-5p.

Cluster Name	Cluster Score	GO Term	Number of Genes	%	*P-value*[a]
Innate immunity in response to stimulus	7.14	GO:0006952~defense response	41	12.73	4.64E-09
		GO:0006954~inflammatory response	23	7.14	2.77E-05
		GO:0009611~response to wounding	25	7.76	5.58E-03
		GO:0006950~response to stress	62	19.25	2.59E-05
		GO:0050896~response to stimulus	90	27.95	2.13E-02
		GO:0045087~innate immune response	18	5.59	5.71E-07
		GO:0002376~immune system process	46	14.29	7.34E-06
		GO:0006955~immune response	36	11.18	1.23E-05
		GO:0009605~response to external stimulus	41	12.73	4.23E-05
Immune response to bacterium	6.34	GO:0051704~multi-organism process	27	8.39	2.11E-02
		GO:0051707~response to other organism	23	7.14	6.09E-06
		GO:0009607~response to biotic stimulus	24	7.45	9.73E-05
		GO:0009617~response to bacterium	20	6.21	1.72E-06
		GO:0006952~defense response	41	12.73	4.64E-09
		GO:0042742~defense response to bacterium	14	4.35	2.86E-05
		GO:0032496~response to lipopolysaccharide	8	2.48	2.53E-02
		GO:0002237~response to molecule of bacterial origin	8	2.48	3.89E-02
Positive regulation of metabolic process	4.05	GO:0048522~positive regulation of cellular process	59	18.32	3.07E-03
		GO:0031325~positive regulation of cellular metabolic process	31	9.63	3.24E-02
		GO:0010604~positive regulation of macromolecule metabolic process	32	9.94	1.67E-02
		GO:0048518~positive regulation of biological process	69	21.43	5.39E-05
		GO:0009893~positive regulation of metabolic process	33	10.25	2.18E-02
		GO:0010557~positive regulation of macromolecule biosynthetic process	25	7.76	3.87E-02
Immune Cell Activation	3.69	GO:0050776~regulation of immune response	15	4.66	8.96E-03
		GO:0001817~regulation of cytokine production	12	3.73	2.58E-02
		GO:0001775~cell activation	17	5.28	9.19E-03
		GO:0045321~leukocyte activation	14	4.35	2.50E-02
Regulation of immune response	3.56	GO:0050776~regulation of immune response	15	4.66	8.96E-03
		GO:0048583~regulation of response to stimulus	22	6.83	1.32E-02
		GO:0002684~positive regulation of immune system process	14	4.35	2.43E-02
		GO:0048584~positive regulation of response to stimulus	14	4.35	2.37E-02
		GO:0002682~regulation of immune system process	18	5.59	3.30E-02
Regulation of macromolecules and cellular biosinthetyc process	3.55	GO:0060255~regulation of macromolecule metabolic process	87	27.02	1.23E-02
		GO:0010556~regulation of macromolecule biosynthetic process	77	23.91	1.61E-02
		GO:0019222~regulation of metabolic process	93	28.88	1.72E-02
		GO:0031323~regulation of cellular metabolic process	89	27.64	2.23E-02

		GO:0031326~regulation of cellular biosynthetic process	78	24.22	2.36E-02
		GO:0009889~regulation of biosynthetic process	78	24.22	2.52E-02
chemotaxis	3.49	GO:0042330~taxis	12	3.73	1.49E-02
		GO:0006935~chemotaxis	12	3.73	1.49E-02
		GO:0007610~behavior	20	6.21	4.07E-02
Regulation of cytokine production	3.45	GO:0032655~regulation of interleukin-12 production	5	1.55	1.67E-02
		GO:0042035~regulation of cytokine biosynthetic process	8	2.48	2.35E-02
		GO:0001817~regulation of cytokine production	12	3.73	2.58E-02
Developmental process	3.28	GO:0009653~anatomical structure morphogenesis	41	12.73	1.28E-02
		GO:0032502~developmental process	82	25.47	2.43E-02
		GO:0048856~anatomical structure development	68	21.12	3.29E-02
		GO:0007275~multicellular organismal development	74	22.98	4.88E-02

[a] Adjusted *P-value* by Benjamini–Hochberg method [53].

Table 4. Functional annotation cluster for candidate mRNAs targets for miR-142-3p.

Cluster Name	Cluster score	GO Term	Number of genes	%	*P-value*[a]
Regulation of transcription	4.72	GO:0065007~biological regulation	148	50.17	3.53E-02
		GO:0050789~regulation of biological process	145	49.15	1.17E-02
		GO:0050794~regulation of cellular process	142	48.14	6.47E-03
		GO:0019222~regulation of metabolic process	95	32.20	7.95E-04
		GO:0009889~regulation of biosynthetic process	78	26.44	1.57E-03
		GO:0031326~regulation of cellular biosynthetic process	77	26.10	1.87E-03
		GO:0010556~regulation of macromolecule biosynthetic process	74	25.08	2.81E-03
		GO:0080090~regulation of primary metabolic process	88	29.83	6.65E-04
		GO:0060255~regulation of macromolecule metabolic process	86	29.15	7.03E-04
		GO:0051171~regulation of nitrogen compound metabolic process	76	25.76	1.41E-03
Intracellular and protein transport	4.41	GO:0051641~cellular localization	37	12.54	5.66E-04
		GO:0051649~establishment of localization in cell	32	10.85	2.81E-03
		GO:0051179~localization	75	25.42	6.50E-03
		GO:0051234~establishment of localization	67	22.71	1.26E-02
		GO:0006810~transport	66	22.37	1.60E-02
		GO:0046907~intracellular transport	29	9.83	1.14E-03
		GO:0016192~vesicle-mediated transport	26	8.81	1.72E-03
		GO:0045184~establishment of protein localization	30	10.17	2.91E-03
		GO:0015031~protein transport	29	9.83	4.66E-03
		GO:0008104~protein localization	31	10.51	7.91E-03
		GO:0033036~macromolecule localization	34	11.53	1.88E-02
Cell communication	3.52	GO:0010646~regulation of cell communication	35	11.86	6.30E-03
		GO:0009966~regulation of signal transduction	31	10.51	7.86E-03
		GO:0050790~regulation of catalytic activity	29	9.83	1.73E-02
		GO:0065009~regulation of molecular function	31	10.51	2.59E-02
		GO:0042325~regulation of phosphorylation	19	6.44	2.99E-02
		GO:0019220~regulation of phosphate metabolic process	20	6.78	2.04E-02
		GO:0051174~regulation of phosphorus metabolic process	20	6.78	2.04E-02
Positive regulation of metabolic and cellular process	3.44	GO:0048522~positive regulation of cellular process	52	17.63	7.71E-03
		GO:0048518~positive regulation of biological process	54	18.31	1.82E-02
		GO:0009893~positive regulation of metabolic process	29	9.83	4.74E-02
Cell cycle	3.34	GO:0007049~cell cycle	28	9.49	1.19E-02
		GO:0051301~cell division	15	5.08	1.89E-02
		GO:0022403~cell cycle phase	18	6.10	2.32E-02
		GO:0000278~mitotic cell cycle	16	5.42	4.67E-02

[a] Adjusted *P-value* by Benjamini–Hochberg method [53].

Table 5. Correlation of prolificacy-associated miRNA expression by RT-qPCR and microarray expression of mRNA candidate target DE[a].

miRNA	Candidate Target DE Genes	R^{2b}	*P-value*
miR-146a-5p	***PPM1K***	**-0.78**	**0.003**
miR-146a-5p	***TLR1***	**-0.43**	**0.166**
miR-146a-5p	*UBXD8*	-0.35	0.260
miR-146a-5p	*LRRK1*	-0.21	0.520
miR-146a-5p	*CCL8*	0.17	0.600
miR-146a-5p	*CFH*	0.12	0.700
miR-146a-5p	*TMSL8*	0.06	0.860
miR-146a-5p	*BTG2*	0.03	0.920
miR-142-3p	***CPEB2***	**-0.62**	**0.032**
miR-142-3p	*SLC35F5*	-0.26	0.400
miR-142-3p	*BAT1*	0.33	0.290

[a] Expression correlation was measured in ovary of 12 IbxMe F_2 pregnant sows from high and low prolificacy groups for which mRNA expression data were described [27].

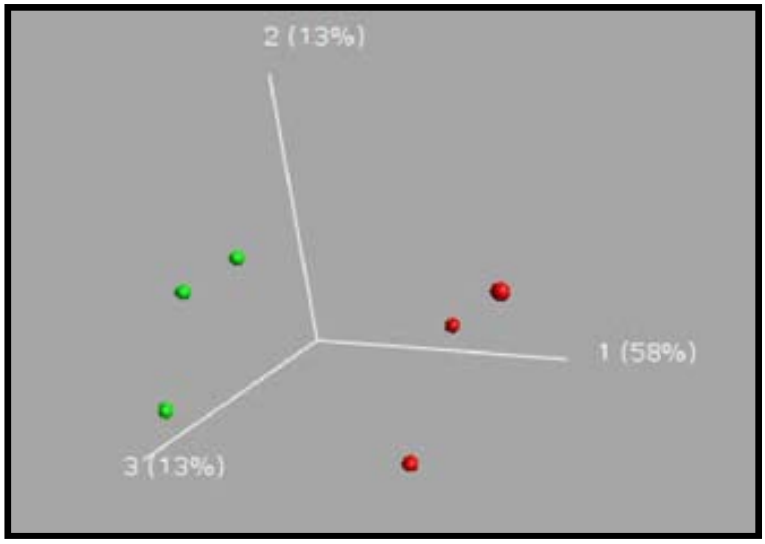
[b] Pearson coefficient.

Table 5. Correlation of prolificacy-associated miRNA expression by RT-qPCR and microarray expression of mRNA candidate target DEa

miRNA		miRNA sequence	Forward Primer Sequence	Reverse Primer Sequence
miR-142-3p	Candidate miRNA	TGTAGTGTTTCCTACTTTATGGA	GCAGTGTAGTGTTTCCTACT	GGTCCAGTTTTTTTTTTTTTTTCCAT
miR-335	Candidate miRNA	TCAAGAGCAATAACGAAAAATG	GCAGTCAAGAGCAATAACGA	AGGTCCAGTTTTTTTTTTTTTTTCATT
miR-146a-5p	Candidate miRNA	TGAGAACTGAATTCCATGGGTT	CAGTGAGAACTGAATTCCATG	GGTCCAGTTTTTTTTTTTTTTTAACC
miR-129-3p	Candidate miRNA	AAGCCCTTACCCCAAAAAGCAT	CAGAAGCCCTTACCCCA	GGTCCAGTTTTTTTTTTTTTTTATGCT
miR-142-5p	Candidate miRNA	CATAAAGTAGAAAGCACTACT	GCAGCATAAAGTAGAAAGCAC	AGGTCCAGTTTTTTTTTTTTTTTAGTAG
miR-146b-5p	Candidate miRNA	TGAGAACTGAATTCCATAGGC	GCAGTGAGAACTGAATTCCA	GTCACGTTTTTTTTTTTTTTTGCCTA
miR-17-3p	Reference miRNA	ACTGCAGTGAAGGCACTTGTAG	GACTGCAGTGAAGGCA	GTCCAGTTTTTTTTTTTTTTTCTACAAG
miR-103	Reference miRNA	AGCAGCATTGTACAGGGCTATGA	AGAGCAGCATTGTACAGG	GGTCCAGTTTTTTTTTTTTTTTCATAG
Let-7a	Reference miRNA	TGAGGTAGTAGGTTGTATAGTT	GCAGTGAGGTAGTAGGTTGT	GGTCCAGTTTTTTTTTTTTTTTAACTATAC

FIGURES

Figure 1. Principal component analysis for the 84 remaining probes after normalisation and filtering process. H is referred to high prolificacy sows whereas L is referred to low prolificacy sows.



ADDITIONAL FILES

Additional Files 1. Microarray expression profile of 84 probes after the filtering steps were applied.

Probe Name	Porcine miRNA Annotation		Mean Expression High Prolificacy	Mean Expression Low Prolificacy	Mean Difference (High-Low)	Fold Change High/Low Prolificacy	Mean Variance	P-value T-test
hsa-miR-146a-5p	ssc-miR-146a-5p	ssc miRNA	8.19	7.53	0.66	1.09	0.182	0.078
hsa-miR-142-3p	ssc-miR-142-3p	ssc miRNA	10.05	9.48	0.58	1.06	0.122	0.013
hsa-miR-142-5p	ssc-miR-142-5p	ssc miRNA	8.84	8.32	0.53	1.06	0.127	0.054
hsa-miR-146b-5p	ssc-miR-146b	ssc miRNA	8.50	7.99	0.51	1.06	0.132	0.118
hsa-miR-452	ssc-miR-452-5p	new pig miRNAs	8.47	8.04	0.43	1.05	0.122	0.194
hsa-miR-223	ssc-miR-223-3p	new pig miRNAs	8.69	8.27	0.42	1.05	0.090	0.098
hsa-miR-144*	ssc-miR-144-5p	new pig miRNAs	6.90	6.49	0.41	1.06	0.190	0.333
hsa-miR-106a	ssc-miR-106a	ssc miRNA	10.03	9.68	0.35	1.04	0.058	0.074
hsa-miR-20a	ssc-miR-20a	ssc miRNA	9.73	9.39	0.34	1.04	0.063	0.117
46879	ssc-miR-452-5p	new pig miRNAs	7.32	6.98	0.34	1.05	0.068	0.178
hsv1-miR-H5	no annotation		9.29	8.95	0.34	1.04	0.104	0.265
hsa-miR-16	ssc-miR-16	ssc miRNA	12.39	12.06	0.33	1.03	0.042	0.041
hsa-miR-221	ssc-miR-221	ssc miRNA	8.53	8.20	0.33	1.04	0.050	0.067
hsa-miR-222	ssc-miR-222	ssc miRNA	9.11	8.79	0.33	1.04	0.067	0.154
hsa-miR-93	ssc-miR-93-5p	new pig miRNAs	9.32	9.00	0.32	1.04	0.035	0.009
hsa-miR-15a	ssc-miR-15a	ssc miRNA	10.83	10.52	0.31	1.03	0.042	0.087
SNORD38B-5	no annotation		9.44	9.13	0.31	1.03	0.033	0.014
mmu-miR-20b	ssc-miR-20b-5p	new pig miRNAs	8.62	8.31	0.30	1.04	0.046	0.073
hsa-miR-132	ssc-miR-132-3p	new pig miRNAs	8.44	8.15	0.30	1.04	0.045	0.130
hsa-miR-18b	ssc-miR-18b	ssc miRNA	7.22	6.93	0.29	1.04	0.033	0.023
hsa-miR-129-3p	ssc-miR-129a	ssc miRNA	6.61	6.32	0.29	1.05	0.031	0.021
hsa-miR-212/mmu-miR-	ssc-miR-212-3p	new pig miRNAs	8.44	8.15	0.28	1.03	0.042	0.132

212/rno-miR-212								
146209	no annotation		10.53	10.25	0.28	1.03	0.037	0.054
hsa-miRPlus-A1070,hsa-miRPlus-D1058	no annotation		7.73	7.45	0.28	1.04	0.031	0.062
146220	no annotation		9.49	9.21	0.28	1.03	0.058	0.195
hsa-miR-27a	ssc-mir-27a	ssc miRNA	10.34	10.07	0.28	1.03	0.037	0.078
hsa-let-7f	ssc-let-7f	ssc miRNA	10.57	10.30	0.28	1.03	0.034	0.056
SNORD65-hsa-mmu	no annotation		7.85	7.58	0.27	1.04	0.033	0.080
kshv-miR-K12-3	no annotation		7.80	7.53	0.27	1.04	0.098	0.358
rno-miR-743b	no annotation		7.57	7.31	0.27	1.04	0.028	0.022
146201	no annotation		7.21	6.95	0.27	1.04	0.042	0.122
hsa-miR-17	ssc-miR-17	ssc miRNA	10.03	9.78	0.26	1.03	0.045	0.152
hsa-miR-129*	ssc-miR-129b-3p	new pig miRNAs	10.03	9.78	0.26	1.03	0.067	0.267
hsa-miR-30a	ssc-miR-30a-5p	ssc miRNA	9.54	9.28	0.26	1.03	0.029	0.048
hsa-miRPlus-F1215	no annotation		7.76	7.50	0.25	1.03	0.036	0.101
hsa-miR-342-3p	ssc-miR-342	ssc miRNA	7.38	7.13	0.25	1.03	0.033	0.086
hsa-miR-1274a	Lys tRNA		11.78	11.54	0.24	1.02	0.061	0.301
hsa-miR-18b	ssc-miR-18b	ssc miRNA	7.99	7.75	0.24	1.03	0.040	0.204
hsa-miR-203	no annotation		7.30	7.06	0.24	1.03	0.020	0.031
hsa-miR-92a	ssc-miR-92a	ssc miRNA	9.20	8.97	0.24	1.03	0.038	0.204
hsa-miR-551b*	no annotation		7.07	6.84	0.23	1.03	0.033	0.121
hsa-miR-21	ssc-miR-21	ssc miRNA	13.67	13.44	0.23	1.02	0.026	0.108
hsa-miR-557	no annotation		7.83	7.60	0.23	1.03	0.089	0.400
hsa-miR-132*	ssc-miR-132-5p	new pig miRNAs	8.05	7.82	0.23	1.03	0.020	0.060
hsa-let-7i	ssc-let-7i	ssc miRNA	11.29	11.06	0.23	1.02	0.023	0.074
hsa-miR-665	ssc-miR-665-3p	new pig miRNAs	7.30	7.07	0.23	1.03	0.236	0.627
hsa-miRPlus-F1074	no annotation		8.22	7.99	0.22	1.03	0.022	0.059
hsa-miR-	ssc-miR-20b-5p	new pig	8.01	7.79	0.22	1.03	0.020	0.075

20b/mmu-miR-20b/rno-miR-20b-5p		miRNAs						
hsa-miR-199a-3p	ssc-miR-199a-3p	ssc miRNA	11.05	10.82	0.22	1.02	0.162	0.581
hsa-miRPlus-A1015	no annotation		8.71	8.49	0.22	1.03	0.026	0.091
mmu-miR-106a	ssc-miR-106a	ssc miRNA	9.49	9.27	0.22	1.02	0.025	0.087
mmu-miR-582-3p	ssc-miR-582-3p	new pig miRNAs	11.04	10.82	0.22	1.02	0.029	0.156
hsa-miR-15b	ssc-miR-15b	ssc miRNA	10.20	9.98	0.22	1.02	0.024	0.075
hsa-miR-21	ssc-miR-21	ssc miRNA	10.00	9.78	0.22	1.02	0.032	0.151
hsa-miR-18b	ssc-miR-18b	ssc miRNA	8.12	7.91	0.21	1.03	0.038	0.269
hsa-miR-1248	HBI-61 or SNORA81		7.45	7.23	0.21	1.03	0.030	0.141
hsa-miR-28-5p	ssc-miR-28-5p	ssc miRNA	7.85	7.64	0.21	1.03	0.041	0.257
hsa-miR-92b*	ssc-miR-92b-5p	ssc miRNA	7.28	7.06	0.21	1.03	0.062	0.371
hsa-miRPlus-F1181	no annotation		7.48	7.27	0.21	1.03	0.041	0.266
hsa-miR-17/mmu-miR-17/rno-miR-17-5p/rno-miR-17	ssc-miR-17	ssc miRNA	9.79	9.59	0.21	1.02	0.027	0.140
hsa-miR-143	ssc-miR-143	ssc miRNA	13.86	13.65	0.21	1.02	0.058	0.361
hsa-miR-155	ssc-miR-155	ssc miRNA	7.72	7.92	-0.20	0.98	0.221	0.663
hsa-miR-125b	ssc-miR-125b	ssc miRNA	13.00	13.20	-0.20	0.98	0.176	0.639
ath-miR771	no annotation		10.11	10.31	-0.20	0.98	0.031	0.208
hsa-miR-675	no annotation		8.40	8.61	-0.20	0.98	0.049	0.321
146147	no annotation		9.55	9.77	-0.22	0.98	0.097	0.479
hsa-miR-1469	no annotation		9.61	9.84	-0.23	0.98	0.040	0.179
hsa-miRPlus-E1151	no annotation		8.43	8.67	-0.24	0.97	0.052	0.284
145997	no annotation		9.84	10.09	-0.25	0.98	0.029	0.101
mmu-miR-	ssc-miR-711-3p	new pig	6.96	7.23	-0.27	0.96	0.030	0.031

711		miRNAs						
hsa-miR-135a/mmu-miR-135a/rno-miR-135a	ssc-miR-135-5p	ssc miRNA	7.86	8.15	-0.29	0.96	0.082	0.263
hsa-miR-583	no annotation		8.99	9.28	-0.29	0.97	0.095	0.307
mmu-miR-710	mouse/rat specific miRNA		7.11	7.43	-0.32	0.96	0.040	0.027
hsa-miR-711	ssc-miR-711-3p	new pig miRNAs	6.54	6.87	-0.33	0.95	0.048	0.079
mmu-miR-881*	Mouse specific miRNA		6.62	6.97	-0.35	0.95	0.106	0.228
146141	no annotation		9.39	9.77	-0.38	0.96	0.052	0.017
hsa-miR-135a	ssc-miR-135-5p	ssc miRNA	8.00	8.38	-0.38	0.95	0.111	0.200
mmu-miR-685	RNase P RNA		7.73	8.16	-0.44	0.95	0.098	0.140
hsa-miR-1290	U2 RNA like		7.56	8.14	-0.58	0.93	0.187	0.163
hsa-miR-335	ssc-miR-335	ssc miRNA	10.47	11.05	-0.58	0.95	0.250	0.181
mmu-miR-2134	rRNA		6.65	7.26	-0.62	0.92	0.176	0.109
mmu-miR-882	part of large cluster of imprinted genes		7.79	8.42	-0.64	0.92	0.170	0.081
hsa-miR-1246	U2 RNA like		10.31	11.20	-0.89	0.92	0.506	0.200
mmu-miR-2137	high resemblence to rRNA (LSU)		8.60	9.77	-1.17	0.88	0.516	0.017

Additional Files 2. *In silico* mRNA candidate targets predicted for miR-146a-5p and miR-142-3p.

miRNA	Target Database	Confidence	Gene Symbol
miR-142-3p	Ingenuity Expert Findings	Experimentally Observed	*BCL2L1*
miR-142-3p	Ingenuity Expert Findings,TargetScan Human	Experimentally Observed,High (predicted)	*BCLAF1, LIFR*
miR-142-3p	TargetScan Human	High (predicted)	*ABL2,ACBD5,ACVR2A,ADAMTS3,AFF1,AFF2,AKT1S1,AMOTL1,ANK3,ANKRD11,ANKRD46,ANKS1A,APC,ARHGEF12,ARHGEF2,ARID5B,ARL1,ARL15,ARNTL,ASH1L,ATF7IP,ATG16L1,ATP2A2,BACH1,BACH2,BAZ1A,BHLHB9,BNC2,BOD1,BRWD3,BTBD7,C10orf18,C13orf33,C16orf70,C18orf25,C1orf9,C20orf194,C4orf39,C5orf24,C9orf72,CCDC165,CCDC6,CCDC88A,CCNJ,CCNT2,CD1D,CD84,CDC25C,CEP192,CFL2,CHRNE,CIITA,CLCN5,CLDN12,CLIC4,CLTA,COG4,COL24A1,COPG,COPS7A,CPEB2,CREB5,CRK,CRTAM,CTNND1,CTTN,CUL5,CXADR,DCUN1D4,DDX39B,DIRC2,DMTF1,DSCR6,DUXA,DYNC1LI2,EDEM3,EHF,EML4,EPM2AIP1,EPN1,ERC1,ERG,FAM114A1,FBXO21,FBXO3,FBXO45,FGF23,FKBP1A,FMNL2,FNDC3A,FNTB,FOXD4L4,FOXO1,FOXO4,GAB1,GFI1,GHR,GNAQ,GNB2,GOLGA1,GPCPD1,GPR85,GTF2A1,HAUS3,HEATR5A,HECTD1,HGS,HMGA2,HMGB1,HSPA1A/HSPA1B,IER3,INPP5A,INPP5F,IRAK1,ITGAV,ITGB8,ITPKB,ITPR3,JMJD1C,KAT2B,KDELR2,KDM6A,KLF13,KLHDC1,LCOR,LHCGR,LLGL2,LMO3,LRP1B,LRRC1,LRRC32,LRRC3B,LRRC59,MANBAL,MAP3K11,MARCH1,MARCKS,MARK3,MBD6,MGAT4A,MGLL,MLL,MMGT1,MNS1,MOBKL3,MON2,MORF4L2,MTCH1,MTMR9,MYH10,MYH9,MYLK,MYO5A,MYST2,MZT1,NK23,nR1D2,NR2C1,NR2F6,NR3C1,NUDT11,PAPD7,PARD6B,PATZ1,PCGF3,PCSK1,PDE4B,PDE8A,PGM1,PGRMC2,PICALM,PIK3R6,PMAIP1,PPFIA1,PPP1R10,PPP1R2,PPP3CA,PPP3R1,PRLR,PROM1,PRRG1,PSME4,PTPN23,PUM1,PURB,RAB11FIP2,RAB1A,RAB2A,RAB3A,RAB40C,RAC1,RAP1GAP2,RARG,RBM27,RBM47,REPS2,RERE,RFWD3,RGL2,RHEB,RHOBTB3,RICTOR,RLF,RNF38,ROCK2,ROD1,RPE,RPRD1A,S1PR3,SAMD12,SDC4,SH2B1,SH3GLB1,SIK1,SIK2,SLC17A5,SLC1A3,SLC25A22,SLC25A44,SLC30A7,SLC35F5,SLC37A3,SLC38A4,SLC39A10,SLCO4C1,SMG1,SMR3A,SNX18,SOCS6,SOX11,SP8,SPNS1,SRSF12,SS18,STAM,STAU1,STRN3,STX12,SYPL1,TAB2,TAGAP,TAOK1,TARDBP,TBC1D2B,TBL1X,TEAD1,TESK2,TET2,TET3,TEX2,TFG,TGFBR1,TIPARP,TIRAP,TKTL2,TMEM110,TMEM115,TMEM200B,TMEM55B,TMEM59,TMTC3,TNKS,TNRC18,TP53INP2,TSEN34,TWF1,UNKL,USP33,USP6NL,USP9X,UTRN,UTY,VAMP3,VPS24,WASL,WDR5B,WHAMM,WIPF2,WIZ,XKRY/XKRY2,XPO1,ZBTB10,ZBTB41,ZCCHC14,ZCCHC24,ZEB1,ZEB2,ZFP91,ZFYVE20,ZMYND8,ZNF217,ZNF395,ZNF440/ZNF808,ZNF479,ZNF618,ZNF676,*

			ZNF682,ZNF827 and *ZNF831*
miR-142-3p	TargetScan Human,miRecords	Experimentally Observed,High (predicted)	*ADCY9*
miR-146a-5p	Ingenuity Expert Findings	Experimentally Observed	*C8A,CAMP,CCR3,CD1D,CD40,CHUK,CRP,DMBT1,IFNA1/IFNA13,IFNB1,IL10,IL12RB2,IL1F10,IL1R1,IL1RAP,IL1RAPL2,IL1RL2,IL36A,IL36B,IL36G,IL36RN,IL37,IL8,LALBA,LBP,LTF,NOS2,PGLYRP1,PGLYRP2,PTAFR,S100A12,SFTPD,TLR1,TLR10,TLR4* and *TLR9*
miR-146a-5p	Ingenuity Expert Findings,miRecords	Experimentally Observed	CFH
miR-146a-5p	Ingenuity Expert Findings,TargetScan Human	Experimentally Observed,High (predicted)	TRAF6
miR-146a-5p	miRecords	Experimentally Observed	*ATOH8,BLMH,CCL8,CCNA2,COL13A1,CXCR4,FADD,IRAK2,IRF5,KIF22,LTB,MCM10,MCPH1,MR1,NFIX,PA2G4,PDIK1L,PEX11G,PLEKHA4,POLE2,PRR15,RAD54L,SDCBP2,STAT1,TMSB15A,TRIM14* and *VWCE*
miR-146a-5p	TargetScan Human	High (predicted)	*ABL2,ACER3,AGMAT,AMPH,AP3M2,APPL1,ARF6,ARL10,ARL8A,ARPC5,ATG12,B3GNT5,BAG1,BCL11A,BCORL1,BEND4,BHLHE41,BIVM,BNC1,BRK1,BTG2,BZW1,C12orf36,C2orf54,C5orf4,C5orf46,C6orf97,C9orf66,C9orf72,CADM2,CARD10,CASK,CASP7,CBARA1,CCBP2,CCDC6,CCDC89,CCK,CCL5,CD80,CD96,CDC42BPA,CDON,CELF2,CLASP2,CNTF,CNTFR,COL4A3,COL9A1,COPS8,COX15,CPLX4,CPM,CSF1R,CXorf1,CYP3A43,DCAF12,DCAF17,DCN,DCTN5,DDX58,DGKG,DLGAP1,DLGAP2,DNAJC12,DNAL1,DNPEP,DPY19L2,DTNA,EDNRB,EIF4G2,EIF5A2,ERBB4,FAF2,FAM110B,FAM120AOS,FAM26E,FAM65B,FANCM,FBXL3,FBXO4,FBXW2,FLOT2,FRYL,GALNT10,GJC1,GOSR1,GRID1,GRSF1,HNRNPD,IER5L,IGSF1,IL17A,INTS2,ITCH,JAZF1,KBTBD3,KCMF1,KCNIP3,KCNJ16,KCTD15,KDM2B,KDM6B,KIAA1161,KIAA1199,KIAA1804,KIF24,KIR3DX1,KLF7,KPNA6,KRT80,KRTAP13-1,LANCL1,LFNG,LIN28A,LRP2,LRRC15,LRRK1,LRRTM2,MARCH6,MED1,MED20,MLL2,MOBKL1A,MRS2,MS4A1,MTAP,MYBL1,MYO6,MYT1,NEB,NF2,NFASC,NOVA1,NOX4,NPAS4,NRAS,NUMB,ONECUT2,OPALIN,OR9Q1,PAQR5,PBX2,PCBP4,PDGFRA,PGK2,PHF20L1,PHKB,PHOX2B,PLSCR4,PM20D2,PMAIP1,POFUT2,PPBP,PPM1K,PPM1M,PPP1R11,PPP2R1B,PRDX4,PRKCE,PRMT3,PSMD3,PTAR1,PTGFRN,RAB8B,RABGAP1,RARB,RASGRP1,RCSD1,REEP5,RFTN2,RHOBTB3,RHOXF2/RHOXF2B,RIMS2,RNASEL,ROBO1,RPL34,RUNX1T1,SCN3B,SEC23IP,SEMA3G,SEPT14,SH3TC2,SIAH2,SIX4,SLC10A3,SLC16A14,SLC19A3,SLC1A1,SLC2A3,SLITRK3,SMAD4,SNX21,SORT1,SP8,SRP72,SRPRB,SRRM4,SRSF12,SRSF6,ST13,ST7L,ST8SIA4,STIL,STRA13,STRBP,STX3,SYN2,SYNJ2,SYNPO2,SYT1,SYT13,SYT14,TAF9B,TANC2,TBRG1,TBX18,TCF*

			21,TDRKH,TEX261,THAP5,TMEM100,TMEM19,TMEM194A,TOR1A,TPD52,TRIT1,TRMT1L,TTC39A,TTPAL,UPP2,USP3,USP47,VPS54,WASF2,WDR47,WWC2,YES1,ZBTB2,ZC3H12B,ZDHHC13,ZFP91-CNTF,ZFYVE1,ZNF117,ZNF138,ZNF148,ZNF229,ZNF253,ZNF257,ZNF331,ZNF354B,ZNF493,ZNF506,ZNF512B,ZNF532,ZNF540,ZNF649,ZNF676,ZNF813,ZNF826 and *ZNRF3*
miR-146a-5p	TargetScan Human	High (predicted),Moderate (predicted)	RGPD5
miR-146a-5p	TargetScan Human,miRecords	Experimentally Observed,High (predicted)	*IRAK1, MMP16, TIMELESS* and *UHRF1*
miR-146a-5p	TargetScan Human,miRecords	Experimentally Observed,Moderate (predicted)	BRCA1, CDKN3, METTL7A and PBLD

4. GENERAL DISCUSSION

4. GENERAL DISCUSSION

Litter size, as many other traits of economical interest in swine industry, is a complex trait which is influenced by other traits like the ovulation rate, the fertilization rate and the embryonic and foetal survival rate (Distl, 2007). The study of this trait is difficult due to the fact that it has a low heritability, long generation intervals and strong environmental influences. Moreover, it is a polygenic trait; thus, the observed phenotype is the result of the effect of several genes that establish complex genetic networks. In a previous work using the same resource population, two QTL for NBA and TNB on SSC13 and SSC17 were identified. In addition, several epistatic QTL were determined showing the complex genetic architecture of prolificacy related traits (Noguera *et al.*, 2009). Although it was evidenced that there were regions in the genome associated with litter size, the genes, which are the real responsible for the genetic variation underlying these genomic regions, remained unknown. Therefore, as a next step in the previous research, the present thesis emerged with the main objective of identifying those genes or genomic regions responsible for the genetic variation associated with these traits in the IbxMe F_2 population. For this, several approaches have been performed to improve our understanding of the genetic architecture related to litter size traits in pigs.

A positional candidate gene approach was performed focusing on the QTL for litter size on SSC13 (Noguera *et al.*, 2009). Four genes (*ITIH1*, *ITIH3*, *ITIH4* and *MUC4*) were selected as candidate genes because they mapped, by comparative linkage mapping between human and pigs, within the confidence interval (CI) for these QTL described by (Noguera *et al.*, 2009); and also because they play a role in reproductive processes. During the gestation, after embryo implantation, all of them participate in the establishment of a thick gycocalyx on the apical uterine surface epithelium to protect the endometrium from proteolytic activity of porcine conceptus and from microbial invasion. Moreover, *ITIH1*, *ITIH3*, *ITIH4* are clustered on SSC13 which reinforce their interest as candidate genes.

The molecular characterisation of porcine *ITIH* gene cluster included the sequencing of the full coding sequences of the three genes. Successfully, we could identify, for the

first time, 3, 9 and 4 SNPs in *ITIH-1*, *-3* and *-4*, respectively. On the other hand, the porcine *MUC4* gene is very large, 31.12 Kb encoding a transcript of 6.4 Kb. Thus, sequencing the full gene sequence as well as the coding region is a demanding task. Alternatively, in porcine *MUC4* gene, two SNPs in intronic regions were reported (Peng *et al.*, 2007) and therefore, we designed a genotyping protocol and validated their segregation in our resource population.

Association studies determined significant effects on NBA for 2 ($C744G_{(\text{that led to Phe248Leu aminoacid change in von Willebrand functional domain (vWA))}}$ and $G2623A_{(Glu875Lys)}$, 4 ($T1332C_{(vWA)}$, $T1341C_{(vWA)}$, $T1362C_{(vWA)}$ and $A1867G_{(Met623Val)}$), 4 ($A826G_{(Ile276Val,\ vWA)}$, $C840T_{(vWA)}$, $G843A_{(vWA)}$ and $A1027G_{(Lys343Glu,\ vWA)}$) and 1 (DQ124298:g.344A>G) SNPs in *ITIH1*, *ITIH3, ITIH4* and *MUC4* genes, respectively. Only DQ124298:g.344A>G SNP in *MUC4* had a significant effect on TNB although it was weaker than the effect on NBA. All favourable effects for these SNPs came from the Meishan breed, as it was expected taking into consideration that Meishan is considered one of the most hyperprolific breeds. The additive effect on NBA was estimated to be about 0.7 for the two *ITIH1* SNPs and the first three *ITIH3* SNPs and 0.4 for the remaining *ITIH3* SNP and for all the *ITIH4* SNPs. Finally, for *MUC4* SNP, the additive substitution effect was estimated to be 0.7 for NBA and 0.6 for TNB.

By considering the *ITIH* SNPs with significant effect on NBA, 2 haplotypes were arranged for *ITIH1* and *ITIH4* while 3 haplotypes were determined for *ITIH3*. To gain more insight about the effect of these genes on NBA, association analyses were performed by using *ITIH* haplotypes and moreover, the QTL effect was included in the analyses to discard false positive association related to nearby prolificacy QTL on the same chromosome 13. For the DQ124298:g.344A>G SNP in *MUC4*, the QTL was also included in the association analysis for the same reason. The effect of *ITIH1* and *ITIH4* haplotypes and DQ124298:g.344A>G SNP in *MUC4* on NBA were confirmed as well as the significant effect of DQ124298:g.344A>G SNP in *MUC4* on TNB. The additive substitution effect for the favourable haplotype, with a Meishan origin, for *ITIH1* and *ITIH4* was estimated to be 0.7 and 0.4, respectively; similar as the previously detected for each individual SNP. This result indicated that the SNPs were in linkage disequilibrium and thus, the resulting additive effect is due to the haplotype instead of

the sum of all SNPs. For *ITIH3* haplotypes, we could only determine suggestive association ($P < 0.1$) between NBA and *ITIH3* haplotypes. However, $ITIH3_A$ and $ITIH3_B$ haplotypes seemed to determine a higher NBA than $ITIH3_C$ haplotype. $ITIH3_A$ and $ITIH3_B$ only differed in the fourth SNP; therefore, it is suggested that the effect on NBA is due to the alleles that these haplotypes shared in the first three. An analysis to know the effect of all *ITIH* gene cluster on NBA would be interesting. However, we could not perform it due to we did not have the statistical power required considering the low sample size available.

The most informative SNP of *ITIH* cluster genes (T886C on *ITIH3* gene) based on polymorphism information content (PIC) (Botstein *et al.*, 1980) as well as the DQ124298:g.344A>G SNP on *MUC4* gene were used to perform a QTL scan on SSC13 in order to refine the prolificacy QTL on SSC13 (Figure 4.1).

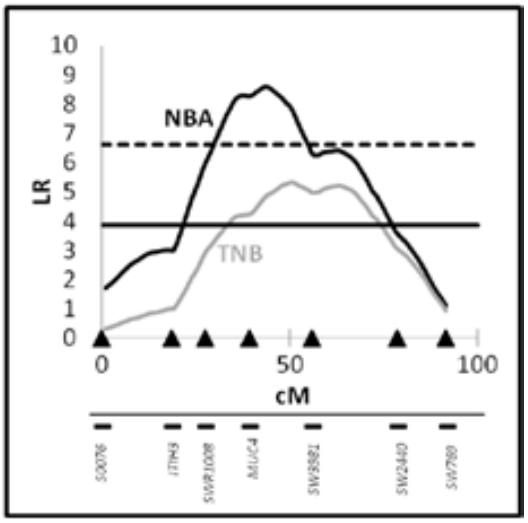


Figure 4.1. QTL profiles on SSC13 for TNB and NBA showing the likelihood ration test statistic. The horizontal line set at 3.84 and 6.63 show the 0.05 and the 0.01 significance levels, respectively.

Unfortunately, *ITIH1* (~ 18 cM) and *ITIH4* (~ 19.1 cM) were not located within the CI of QTL for NBA (22 - 77 cM) on SSC13 and, therefore, they were discarded to be the causal mutations. However, the significant association found for these genes suggested that they had a slight effect on NBA independent of the QTL described by

(Noguera *et al.*, 2009). These effects on NBA could be explained because of the linkage disequilibrium between polymorphisms and the causal mutation. Another possibility can be that the genomic region where these genes are located harbour a QTL with small effect on NBA trait which cannot be determined through QTL analyses due to the small population size and the low density linkage map used. At transcriptional level, no differences in gene expression could be found regarding prolificacy levels. However, the expression of *ITIH3* was differentially expressed between the two samples of uterus analysed (apical (proximal to the ovary) and basal (close to the cervix)) showing the uterine region-dependent expression of this gene. Pigs have a bipartite uterus with a relatively long cornu and short corpus and cervix. Previous studies indicated uterine-dependent differences functions related to autonomic innervation and changes of spontaneous contractility and responsiveness to bioactive agents like acetylcholine, epinephrine, oxytocin and 5-hydroxytryptamine (Cao *et al.*, 2005). As *ITIH3* participate in the establishment of a thick gycocalyx on the apical uterine surface epithelium, the higher *ITIH3* gene expression in apical uterus than in basal uterus could be related to uterine contractility. Taken the expression levels of *ITIH* gene cluster in both uterus regions, we could determine the expression pattern for these genes. *ITIH1* and *ITIH4* are positively correlated whereas it seems that *ITIH3* is regulated independent of the other *ITIH* genes. However, the regulation of *ITIH* gene cluster is not well understood.

On the other hand, *MUC4* (39.1 cM) was successfully mapped within the CI of prolificacy QTL (22 - 77 and 33 - 74 cM for NBA and TNB, respectively). Furthermore, association studies showed that DQ124298:g.344A>G SNP in *MUC4* explained better the QTL effect on NBA and TNB than the QTL itself. Then, DQ124298:g.344A>G SNP in *MUC4* could be hypothesised to be the causal mutations for prolificacy QTL. However, the location of this SNP within an intronic region without any effect on splicing mechanisms, suggested that instead of being the causal mutation it could be in linkage disequilibrium with it. Gene expression analysis of *MUC4* gene showed that different levels of *MUC4* transcript were associated with prolificacy levels. This result suggested that *MUC4* gene sequence could contain mutations that would affect the *MUC4* gene expression which could be the causal mutation. Then, it would be interesting to sequence the whole gene including the regulatory regions in order to identify new mutations that could explain the difference in gene expression.

Candidate gene approach is a very powerful statistical tool to assess association between polymorphisms and traits of interest. However, this approach can only be done for a small number of genes and usually the results obtained are not very conclusive. In this sense, association studies between candidate genes and quantitative traits have been criticized mainly due to the low replication of the results (Georges, 2007). Moreover, when significant associations are detected, usually they only explain a low fraction of the genetic variability. All these limitations are reflected in the fact that only one successful study identified the causal mutation in pigs (Van Laere *et al.,* 2003). For this reason, in the last decades, the use of high-throughput technologies, such as microarrays, has allowed the analysis of gene expression at the genome-wide level. To unravel the genetic variation underlying complex traits it is necessary to combine the information coming from different methodologies, i.e. QTL detection and global analysis of gene expression. Following this strategy, it could be possible the identification of differentially expressed genes that have a role in the phenotype for which a QTL has been described and, then, become new candidate genes to be analysed.

One of the objectives of this thesis was to determine the genes that have differential expression in IbxMe F_2 sows displaying extreme phenotypes regarding the prolificacy levels, measured as the number of embryos attached to the uterus at 30-32 days of gestation. Firstly, few samples from IbxMe F_2 sows with the most extreme values for the NE were selected to be used for the analysis at genome-wide level. The low sample size was because we maximised phenotypical differences against the statistical power. Subsequently, in order to validate the results, the sample size was increased (n = 36) to increase the statistical power with the inconvenience that the phenotypical differences were reduced.

Expression analyses in uterus were performed by microarray technology using the porcine Affymetrix Genechip that allows the evaluation of 20,201 genes simultaneously. As a result, 297 were differentially expressed (Fold Change > 2 and FDR < 0.05) regarding prolificacy levels. In high prolificacy sows, 101 genes were down-regulated whereas 196 genes were up-regulated. It is worth mentioning that DE genes included *MUC4* gene but not the *ITIH* gene cluster, which is in agreement with our

previous analyses. *In silico* functional analyses of the dysregulated genes offered a better understanding of which are the biological processes affecting prolificacy levels for which expression changes have an impact on litter size. It was shown that uterus of high prolificacy sows is better protected from porcine proteolytic conceptuses and from microbial invasion by the up-regulation of genes involved in muscle tissue development processes that might participate in the establishment of a thick gycocalyx on the apical uterine surface epithelium necessary to protect the endometrium (Ferrell *et al.*, 2003,Keys and King, 1990). Furthermore, down-regulation of genes that participate in cell junction organization and adhesion would avoid the access of different substrates to the cell surface including microbes benefiting by this way the protection of the uterus. Importantly, these endometrial responses evoked by the developing pig conceptus during uterine attachment resemble the acute phase response induced during generalized tissue inflammation (Geisert and Yelich, 1997,Salier *et al.*, 1996)(Geisert and Yelich, 1997,Salier *et al.*, 1996) that would explain the up-regulation of genes related to the inflammatory response in high prolificacy sows. Surprisingly, in high prolificacy sows we observed a down-regulation of genes involved in muscle and circulatory system process, protein transport and in muscle tissue development although they have been reported to be beneficial for the embryo survival (Croy *et al.*, 2009,Tayade *et al.*, 2007). In front of this result, we hypothesized that the overexpression of these genes would be harmful for the embryonic development. Another possibility could be that genes involved in these biological processes have different functions which have not yet been reported. It has to be considered that the annotation of genes came from human annotation and thus, although the homology with pigs is high there might be genes without homology. Thus, differences in gene function between human and pig could be another explanation.

These differentially expressed genes can be considered as functional candidate genes for the prolificacy QTL described by (Noguera *et al.*, 2009). To relate the results obtained by a QTL detection approach and a global gene expression approach, one might be conscious that in a QTL approach the causal mutation is assumed to be at the genetic level, that is a polymorphism within the nucleotide sequence, whereas with the microarray approach one can detect differences in the expression level regarding a

particular phenotype but the causal mutation that determines such differences at the expression level are not known. They could be mutations in the promoter regions of the genes, but also in regulatory elements such as microRNAs.

As described by (Noguera *et al.*, 2009), epistasis has a large effect on prolificacy related traits in the IbxMe F_2 population. To decipher which are the gene interactions underlying the epistatic QTL identified in the IbxMe F_2 population, we decided to characterise the microRNAome in uterus and ovary samples, with the aim of identifying target genes that have shown differential expression between high and low prolificacy sows.

To characterise the microRNAome in the uterus and in the ovary two different approaches were used: high throughput sequencing (HTS) (for uterus) and miRNA microarray (for ovary). Although both can lead to the same results there are some differences. Samples preparation and results analysis through microarray technology is easier and cheaper than in high throughput sequencing technology. However, the number of miRNAs that can be evaluated are fixed in microarray and due to the fact that specific porcine chip are not available, only miRNAs conserved among mammals can be tested. Conversely, as an advantage in HTS, all miRNAs expressed in a tissue can be tested but it is necessary a good depth sequencing when differential expression analysis are wanted to be performed.

The microRNAome was characterised in the uterus by using the same samples as used previously in gene expression analysis. For this, high throughput sequencing methodology was applied. A total of 172,901 sequences from 15 to 30 nucleotides were obtained. Interestingly, the majority of these sequences were annotated as miRNAs (95.3%) evidencing the high efficiency of the protocol used for miRNA detection. Unfortunately, the depth sequencing obtained was insufficient to perform an analysis of differential expression. This lack is due to the small difference in miRNA expression level regarding prolificacy levels. Therefore, this technique could only be used to describe the miRNA expression profile. By this way, miR-125b-5p, miR-200c-3p, miR-23b-3p, miR-23a-3p and miR-99a-5p were identified as the most abundant miRNAs in both prolificacy groups representing about the 70% of the total miRNAs.

Among the literature, several miRNAs (i.e. miR-139-5p, miR-150-5p, miR-27a-3p, miR-20-5p, miR-21-5p, miR101-3p, miR-122-5p, let-7d-5p and miR-199b-3p) (Lei *et al.*, 2011,Abd El Naby *et al.*, 2011,Carletti and Christenson, 2009) were described as important for reproductive processes. Differential expression regarding prolificacy levels of four miRNAs (miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p) (prolificacy-related miRNAs) could be determined but we cannot discard the presence of more differentially miRNAs. Thus, it is planned to resequence the small RNA libraries to obtain the depth sequencing necessary to perform differential expression analysis and therefore, identifying all miRNAs differentially expressed.

miRNAs are regulatory molecules that target several mRNA regulating their expression levels. Prolificacy-related miRNAs in uterus could target several predicted mRNAs. Interestingly, among these, 94 of these predicted targets were found differentially expressed between the high and low prolificacy IbxMe F_2 sows. Correlation expression analysis between miRNAs and mRNAs suggested 40 significant correlations. The significant negative correlation suggested the direct effect of miRNA on their target by downregulating the mRNA expression level in concordance with the traditional role associated to miRNAs. However, the number of significant positive correlation was higher than the negative. Recently, it has been described that miRNAs can also modulate positively the gene expression (Vasudevan, 2011,Vasudevan *et al.*, 2007). This would explain the positive correlations found although another possibility could be that the time of measure the gene expression did not reflect the gene expression downregulation considering that miRNAs fine tune the gene expression and thus, the decrease of it is slightly over the time. Overall, it is suggested that prolificacy miRNAs could have a direct effect, positively or negatively, on the gene expression level and they could be new post-transcriptional mechanisms. Nevertheless, we cannot discard the direct effect of prolificacy miRNAs on the other genes but it could not be determined. This lack in the correlation detection could be also associated to the fact that miRNAs act as fine-tunning mechanisms and the increase/decrease level of mRNA expression could not be detected at time of measure. Overall, these 94 differentially expressed mRNAs that can be regulated by prolificacy miRNAs were considered good candidate genes for prolificacy in the IbxMe population. Moreover, to increase the

efficiency in the candidate gene detection, the QTL information was added. In this sense, 32 genes including *MUC4* gene were located within confidence interval for significant QTL described. Overall, these genes are differentially expressed regarding prolificacy which expression levels can be regulated by miRNAs and moreover, they are located within confidence intervals for prolificacy QTL. As different criteria for the selection of these genes have been applied, the success probability to identify the causal mutation in these genes increased. An important point to be considered is that putative targets have been predicted by using *in silico* predictions. By using this methodology, several false positive results could be obtained. Importantly, only for *MUC4* gene evidence to be controlled by miR-150-5p has been reported (Srivastava *et al.*, 2011).

A previous work performed by our group determined ovarian differential expression regarding prolificacy levels (Fernandez-Rodriguez *et al.*, 2011) by using the same approach as used in the present thesis for the analysis of uterine transcriptome. In the present thesis, we have characterised the ovarian microRNAome. For this, instead of using HTS, microarray technology was applied. The miRCURY LNA™ microRNA microarray v.11.0 (Exiqon, Denmark) which contains 1891 capture probes in quadruplicate, targeting miRNAs from human, mouse and rat registered in the miRBase v14.0 was used. Analyses of miRNA microarrays showed that expression differences between prolificacy groups were very low. In this way, only 50 porcine miRNAs showed the major expression differences but any of them passed the statistical threshold of FDR < 0.05 although they separated the high from the low prolificacy sows by principal component analysis. It has to be considered the low sample size used (n=6) and the fact that sows used have similar genetic background because they came from the same F_2 intercross could be the reasons for the low differences in gene expression. The five miRNAs that showed maximal expression differences regarding the prolificacy levels (miR-146a-5p, miR-142-3p, miR-142-5p, miR-146b and miR-335) were selected to be profiled by RT-qPCR to validate the microarray results although the significance threshold was high (*P-Value* < 0.20). Moreover, to confirm the results of RT-qPCR validation, the sample size used was expanded. Finally, we could identify differential expression for miR-146a-5p and miR-

142-2p associated to prolificacy levels. The biological role of these miRNAs based on their putative targets indicated that they are involved mainly in immune system process and in the regulation of processes related with the regulation of the transcription, transport and cell communication and in the regulation of metabolic processes. These miRNAs could regulate the expression levels of 11 targets differentially expressed in ovary but only direct negatively affect could be determined for three genes. Nevertheless, we cannot discard the effect of prolificacy-related miRNAs on the expression levels due the fine-tunnign mechanisms relate with miRNA. Interestingly, among these 11 predicted targets for prolificacy-related miRNAs that were differentially expressed in the ovary associated to prolificacy levels, 4 of them (*LRRK1*, *CCL8*, *CPEB2* and *BAT1*) have been located within CI of epistatic QTL reported by Noguera *et al.*, (2009).

4.1. Impact of the current research

By positional candidate gene approach, *MUC4* has been postulated as an important candidate gene responsible for the prolificacy QTL found on SSC13 by Noguera *et al.*, (2009) in the same resource population as the one used in the current thesis. On the other hand, we have provided, for the first time, the analysis of the transcriptome and the microRNAome in uterus of pregnant sows regarding prolificacy traits which has provided an overview of the complex regulatory network. Furthermore, the characterisation of ovarian microRNAome was evaluated. Combining the information of the gene expression regulatory network could be used as a tool to identify potential new candidate genes.

Overall, our results, by using different approaches, suggested *MUC4* gene to be a good candidate gene for prolificacy QTL on SSC13. Further work should be performed to further characterise this gene. It would be interesting to analyse the sequence of this gene including the regulatory regions in order to identify mutation within it that could explain the differences in gene expression regarding prolificacy levels. Moreover, functional analysis should be performed in order to understand the post-transcriptional mechanism associated with miR-150-5p.

By the analysis of transcriptome and microRNAome, we could better understand the regulatory networks. miRNAs have a main role associated with prolificacy levels although the differences in fold changes between high and low prolificacy sows is very small. However, these small differences could be translated to a large phenotypic variation due to the fact that they can target hundreds of different targets (Calin and Croce, 2006). The importance of the immune system in prolificacy levels has been evidenced by the fact that many differentially expressed genes in uterus as well as in ovaries are related with it. Moreover, miR-146a-5p, differentially expressed in ovary also has an important role related to immune system. Combing gene and miRNA expression with QTL location, several candidate genes have been proposed to be good candidate genes for prolificacy QTL. However, as target prediction has been predicted *in silico*, functional analysis would be of great interest in order to confirm miRNA targets and to validate the new post-transcriptional molecular mechanisms.

Nowadays, several tools that allow the study at genome wide level have developed. Combining different strategies is a powerful tool to increase the probability of identifying good candidate genes. For example, the high density SNPs chips designed in pigs would be of great interest in order to perform association analysis for SNPs within candidate genes proposed. Moreover, it could be very useful in order to decrease the confidence intervals of prolificacy QTL that would decrease the number of genes underlying these QTL and thus, it would become easier to identify the causal mutation for prolificacy QTL.

5. CONCLUSIONS

1. The characterisation of porcine *Inter-alpha-trypsin inhibitor heavy chain* (*ITIH*) -*1*, -*3* and -*4* genes has revealed the existence of 10 SNPs within these genes that has been associated with the number of piglets born alive in the IbxMe F_2 population. The *ITIH* gene cluster has an effect on the number of piglets born alive, independent of the QTL for the number of piglets born alive located downstream on the porcine chromosome 13, explaining a proportion of 3% of the phenotypic variance.

2. A polymorphism (DQ124298:g.344A>G) in the porcine *MUC4* gene, located within the prolificacy QTL described by (Noguera *et al.*, 2009), is associated with the number of piglets born alive and with the total number of piglets born determining an increase of 0.7 and 0.6 piglets, respectively. Moreover, *MUC4* gene in the uterus is two-fold increase in high prolificacy sows than in low prolificacy IbxMe F_2 sows at 30-32 days of gestation.

3. The gene expression analyses in the uterus of pregnant IbxMe F_2 sows displaying extreme phenotypes for the number of embryos attached to the uterus on day 30-32 of gestation have identified 297 genes differentially expressed. In high prolificacy sows, the 101 up-regulated genes are involved in the inflammatory response to stimulus and muscle tissue development; while the 196 downregulated genes participate in muscle tissue development, cell junction organization and adhesion, biological regulation, muscle and circulatory system processes, and transport.

4. The miRNAs expression profile in the uterus of pregnant IbxMe F_2 sows has showed that miR-125b-5p, miR-200c-3p, miR-23b-3p, miR-23a-3p and miR-99a-5p are the most abundant miRNAs representing about the 70% of the total miRNAs. Regarding prolificacy levels, miR-139-5p, miR-150-5p, miR-27a-3p and miR-20-5p (prolificacy-related miRNAs) are differentially expressed with small differences in fold change being downregulated in high prolificacy sows.

5. In the uterus, prolificacy-related miRNAs could target 94 out of the 297 genes differentially expressed genes associated to prolificacy levels. Thirty-two predicted target genes are located within the confidence interval of the QTL for

the number of piglets born alive and with the total number of piglets born described in (Noguera *et al.*, 2009) and are, therefore, proposed to be good candidate genes for prolificacy QTL. Among the candidate genes to be considered, *MUC4* gene is of great interest to be the responsible for the QTL for litter size on porcine chromosome 13 because it fulfils several criteria: it is located within the QTL confidence interval, its expression level varies regarding the prolificacy level of sows, and it is targeted by the prolificacy-related miRNA miR-150-5p.

6. The characterization of ovarian microRNAome has revealed that miR-146a-5p and miR-142-3p (prolificacy-related miRNAs) are differentially expressed in the ovary of pregnant IbxMe F_2 sows regarding prolificacy levels with also small differences in fold change and being upregulated in high prolificacy sows. Genes targeted by both prolificacy miRNAs are involved in immune system processes and cellular homeostasis.

7. Four genes (*LRRK1*, *CCL8*, *CPEB2* and *BAT1*), putative targets for prolificacy-related miRNAs in the ovary, are proposed to be good candidate genes responsible for the prolificacy epistatic QTL due to: their expression in the ovary is associated with the prolificacy levels of the sows, they are putative targets for ovarian prolificacy-related miRNAs and they are located within the confidence interval of epistatic QTL for the number of piglets born alive and with the total number of piglets born traits.

8. RT-qPCR with two miRNA-specific DNA primers is a highly specific, sensitive and accurate method for miRNA quantification.

6. REFERENCES

Abd El Naby W.S., Hagos T.H., Hossain M.M., Salilew-Wondim D., Gad A.Y., Rings F., Cinar M.U., Tholen E., Looft C., Schellander K., Hoelker M. & Tesfaye D. (2011) Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. *Zygote (Cambridge, England)*, 1-21.

Alonso-Spilsbury M., Mota-Rojas D., Villanueva-Garcia D., Martinez-Burnes J., Orozco H., Ramirez-Necoechea R., Mayagoitia A.L. & Trujillo M.E. (2005) Perinatal asphyxia pathophysiology in pig and human: a review. *Animal Reproduction Science* 90, 1-30.

Alves E., Ovilo C., Rodriguez M.C. & Silio L. (2003) Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. *Animal Genetics* 34, 319-324.

Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nature reviews.Genetics* 2, 130-138.

Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M. & Hansson I. (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science (New York, N.Y.)* 263, 1771-1774.

Bagga S., Bracht J., Hunter S., Massirer K., Holtz J., Eachus R. & Pasquinelli A.E. (2005) Regulation by let-7 and lin-4 miRNAs Results in Target mRNA Degradation. *Cell* 122, 553-563.

Baley J. & Li J. (2012) MicroRNAs and ovarian function. *Journal of ovarian research* 5, 8.

Balcells I., Cirera S. & Busk P.K (2011) Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnology* 11, 70.

Bartel D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215-233.

Bartel D.P. & Chen C.Z. (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature reviews.Genetics* 5, 396-400.

Bell J.T. & Spector T.D. (2011) A twin approach to unraveling epigenetics. *Trends in genetics : TIG* 27, 116-125.

Bentwich I., Avniel A., Karov Y., Aharonov R., Gilad S., Barad O., Barzilai A., Einat P., Einav U., Meiri E., Sharon E., Spector Y. & Bentwich Z. (2005) Identification of hundreds of conserved and nonconserved human microRNAs. *Nature genetics* 37, 766-770.

Berezikov E. (2011) Evolution of microRNA diversity and regulation in animals. *Nature reviews.Genetics* 12, 846-860.

Bergsma R., Kanis E., Verstegen M.W.A. & Knol E.F. (2008) Genetic parameters and predicted selection results for maternal traits related to lactation efficiency in sows. *Journal of animal science* 86, 1067-1080.

Bidanel, J. (2011). Biology and Genetics of Reproduction. In: M. F. Rothschild and A. Ruvinsky (eds.) The genetics of the pig. p 218-241.

Bidanel, J.P., Caritez, J.C. & Legault, C. (1989). Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. Genet. Sel. Evol. 21, 507–526.

Blasco, A., Bidanel, J.P. and Haley, C.S. (1995) Genetics and neonatal survival. In: Varley, M.A. (ed.) The Neonatal Pig: Development and Survival. CAB International, Wallingford, UK, pp. 17–38.

Bonnet A., Le Cao K.A., Sancristobal M., Benne F., Robert-Granie C., Law-So G., Fabre S., Besse P., De Billy E., Quesnel H., Hatey F. & Tosser-Klopp G. (2008) In vivo gene expression in granulosa cells during pig terminal follicular development. *Reproduction (Cambridge, England)* 136, 211-224.

Bost F., Diarra-Mehrpour M. & Martin J.P. (1998) Inter-alpha-trypsin inhibitor proteoglycan family--a group of proteins binding and stabilizing the extracellular matrix. *European journal of biochemistry / FEBS* 252, 339-346.

Botstein D., White R.L., Skolnick M. & Davis R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32, 314-331.

Braglia S., Ramírez O., Noguera J.L., Tomás A., Ovilo C. & Varona L. (2006) Comparison of genetic models for analysing the effects of a PvuII polymorphism in the oestrogen receptor 1 (ESR1) gene on prolificacy in an Iberian ɔMeishan pig population. *Animal Genetics* 37, 454-458.

Brien, F.D. (1986) a review of the genetic and physiological relationships between growth and reproduction in mammals. *Animal Breeding Abstracts* 54, 975-997.

Busk PK: Method for Quantification of Small RNA Species. 2010, WO/2010/085966.

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. *Journal of Animal Breeding and Genetics* 122, 259-263.

Buske B., Sternstein I. & Brockmann G. (2006) QTL and candidate genes for fecundity in sows. *Animal Reproduction Science* 95, 167-183.

Caetano A.R., Johnson R.K., Ford J.J. & Pomp D. (2004) Microarray profiling for differential gene expression in ovaries and ovarian follicles of pigs selected for increased ovulation rate. *Genetics* 168, 1529-1537.

Calin G.A. & Croce C.M. (2006) MicroRNA-cancer connection: the beginning of a new tale. *Cancer research* 66, 7390-7394.

Canario L., Pere M.C., Tribout T., Thomas F., David C., Gogue J., Herpin P., Bidanel J.P. & Le Dividich J. (2007a) Estimation of genetic trends from 1977 to 1998 of body composition and physiological state of Large White pigs at birth. *Animal : an international journal of animal bioscience* 1, 1409-1413.

Canario L., Rydhmer L., Gogue J., Tribout T. & Bidanel J.P. (2007b) Estimation of genetic trends from 1977 to 1998 for farrowing characteristics in the French Large White breed using frozen semen. *Animal : an international journal of animal bioscience* 1, 929-938.

Canario L., Cantoni E., Le Bihan E., Caritez J.C., Billon Y., Bidanel J.P. & Foulley J.L. (2006) Between-breed variability of stillbirth and its relationship with sow and piglet characteristics. *Journal of animal science* 84, 3185-3196.

Cao J., Yosida M., Kitazawa T. & Taneike T. (2005) Uterine region-dependent differences in responsiveness to prostaglandins in the non-pregnant porcine myometrium. *Prostaglandins & other lipid mediators* 75, 105-122.

Carlborg O. & Haley C.S. (2004) Epistasis: too often neglected in complex trait studies? *Nature reviews.Genetics* 5, 618-625.

Carletti M.Z., Fiedler S.D. & Christenson L.K. (2010) MicroRNA 21 blocks apoptosis in mouse periovulatory granulosa cells. *Biology of reproduction* 83, 286-295.

Carletti M.Z. & Christenson L.K. (2009) MicroRNA in the ovary and female reproductive tract. *Journal of animal science* 87, E29-38.

Carraway K.L., Theodoropoulos G., Kozloski G.A. & Carothers Carraway C.A. (2009) Muc4/MUC4 functions and regulation in cancer. *Future oncology (London, England)* 5, 1631-1640.

Carraway K.L., Perez A., Idris N., Jepson S., Arango M., Komatsu M., Haq B., Price-Schiavi S.A., Zhang J. & Carraway C.A. (2002) Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. *Progress in nucleic acid research and molecular biology* 71, 149-185.

Carraway K.L. & Idris N. (2001) Regulation of sialomucin complex/Muc4 in the female rat reproductive tract. *Biochemical Society transactions* 29, 162-166.

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. *Journal of animal science* 79, 623-633.

Chakrabarty A., Tranguch S., Daikoku T., Jensen K., Furneaux H. & Dey S.K. (2007) MicroRNA regulation of cyclooxygenase-2 during embryo implantation. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15144-15149.

Chang C.Y., Chang H.W., Chen C.M., Lin C.Y., Chen C.P., Lai C.H., Lin W.Y., Liu H.P., Sheu J.J. & Tsai F.J. (2011) MUC4 gene polymorphisms associate with endometriosis development and endometriosis-related infertility. *BMC medicine* 9, 19.

Cheloufi S., Dos Santos C.O., Chong M.M. & Hannon G.J. (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465, 584-589.

Chen C., Ridzon D.A., Broomer A.J., Zhou Z., Lee D.H., Nguyen J.T., Barbisin M., Xu N.L., Mahuvakar V.R., Andersen M.R., Lao K.Q., Livak K.J. & Guegler K.J. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic acids research* 33, e179.

Chen C.C., Chang T. & Su H.Y. (2004) Characterization of porcine leptin receptor polymorphisms and their association with reproduction and production traits. *Animal Biotechnology* 15, 89-102.

Chen, C., Chang, T. and Su, H. (2004a) Genetic polymorphisms in porcine leptin gene and their association with reproduction and production traits. Australian Journal of Agricultural Research 55, 699–704

Chen L., Wert S.E., Hendrix E.M., Russell P.T., Cannon M. & Larsen W.J. (1990) Hyaluronic acid synthesis and gap junction endocytosis are necessary for normal expansion of the cumulus mass. *Molecular reproduction and development* 26, 236-247.

Chen P., Baas T.J., Mabry J.W., Koehler K.J. & Dekkers J.C.M. (2003) Genetic parameters and trends for litter traits in U.S. Yorkshire, Duroc, Hampshire, and Landrace pigs. *Journal of animal science* 81, 46-53.

Chiang H.R., Schoenfeld L.W., Ruby J.G., Auyeung V.C., Spies N., Baek D., Johnston W.K., Russ C., Luo S., Babiarz J.E., Blelloch R., Schroth G.P., Nusbaum C. & Bartel D.P. (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes & development* 24, 992-1009.

Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibé B, Bouix J, Caiment F, Elsen JM, Eychenne F, Larzul C, Laville E, Meish F, Milenkovic D, Tobin J, Charlier C, Georges M. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nature Genetics. 38,813-818.

Christenson R.K., Leymaster K.A. & Young L.D. (1987) Justification of unilateral hysterectomy-ovariectomy as a model to evaluate uterine capacity in swine. *Journal of animal science* 65, 738-744.

Cochrane D.R., Cittelly D.M. & Richer J.K. (2011) Steroid receptors and microRNAs: relationships revealed. *Steroids* 76, 1-10.

Croy B.A., Wessels J.M., Linton N.F., van den Heuvel M., Edwards A.K. & Tayade C. (2009) Cellular and molecular events in early and mid gestation porcine implantation sites: a review. *Society of Reproduction and Fertility supplement* 66, 233-244.

Curtis S.E. (1974) Responses of the piglet to perinatal stressors. *Journal of animal science* 38, 1031-1036.

Dahan O., Gingold H. & Pilpel Y. (2011) Regulatory mechanisms and networks couple the different phases of gene expression. *Trends in Genetics* 27, 316-322.

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. *Livestock Production Science* 72, 185-198.

Dekkers, J. C., Mathur, P.K. & Knol E.F. (2011). Genetics improvement of the pig. In: M. F. Rothschild and A. Ruvinsky (eds.) The genetics of the pig. p 390-425.

Dekkers J.C. & Hospital F. (2002) The use of molecular genetics in the improvement of agricultural populations. *Nature reviews.Genetics* 3, 22-32.

Diederichs S. & Haber D.A. (2007) Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression. *Cell* 131, 1097-1108.

Dickerson G.E. & Hazel L.N (1944) Effectiveness of selection of progeny performance as a supplement to earlier culling of livestock. *Journal of Agricultural Research* 69, 459-476.

Distl O. (2007) Mechanisms of regulation of litter size in pigs on the genome level. *Reproduction in domestic animals = Zuchthygiene* 42 Suppl 2, 10-16.

Driancourt, M.A., Prunier, A., Bidanel, J.P. & Martinat-Botté, F. (1992) HCG induced oestrus and ovulation rate and FSH concentrations in prepuberal gilts from lines differing by their adult ovulation rate. *Animal Reproduction Science* 29, 297–305.

Drogemuller C., Hamann H. & Distl O. (2001) Candidate gene markers for litter size in different German pig lines. *Journal of animal science* 79, 2565-2570.

Du H., Chen J., Cui J., Wang X. & Zhang X. (2009) Polymorphisms on SSC15q21-q26 Containing QTL for reproduction in Swine and its association with litter size. *Genetics and molecular biology* 32, 69-74.

Ducos, A. and Bidanel, J.P. (1996) Genetic correlations between production and reproductive traits measured on-farm, in the Large White and French Landrace pig breeds. *Journées de la Recherche Porcine en France* 28, 15–22.

Enghild J.J., Thogersen I.B., Pizzo S.V. & Salvesen G. (1989) Analysis of inter-alpha-trypsin inhibitor and a novel trypsin inhibitor, pre-alpha-trypsin inhibitor, from human plasma. Polypeptide chain stoichiometry and assembly by glycan. *The Journal of biological chemistry* 264, 15975-15981.

Esteller M. (2011) Non-coding RNAs in human disease. *Nature reviews.Genetics* 12, 861-874.

Fernandez-Rodriguez A., Munoz 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. *Biology of reproduction* 84, 299-307.

Fernandez-Rodriguez A., Munoz 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. *Biology of reproduction* 84, 299-307.

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. *Animal Genetics* 41, 73-80.

Ferraz J.B. & Johnson R.K. (1993) Animal model estimation of genetic parameters and response to selection for litter size and weight, growth, and backfat in closed seedstock populations of large white and Landrace swine. *Journal of animal science* 71, 850-858.

Ferrell A.D., Malayer J.R., Carraway K.L. & Geisert R.D. (2003) Sialomucin complex (Muc4) expression in porcine endometrium during the oestrous cycle and early pregnancy. *Reproduction in domestic animals = Zuchthygiene* 38, 63-65.

Fiedler S.D., Carletti M.Z., Hong X. & Christenson L.K. (2008) Hormonal regulation of MicroRNA expression in periovulatory mouse mural granulosa cells. *Biology of reproduction* 79, 1030-1037.

Geisert R.D., Ashworth M.D. & Malayer J.R. (2003) Expression of inter-alpha-trypsin inhibitor heavy chains in endometrium of cyclic and pregnant gilts. *Reproduction (Cambridge, England)* 126, 621-627.

Geisert R.D. & Yelich J.V. (1997) Regulation of conceptus development and attachment in pigs. *Journal of reproduction and fertility.Supplement* 52, 133-149.

Gentleman R.C., Carey V.J., Bates D.M., Bolstad B., Dettling M., Dudoit S., Ellis B., Gautier L., Ge Y., Gentry J., Hornik K., Hothorn T., Huber W., Iacus S., Irizarry R., Leisch

F., Li C., Maechler M., Rossini A.J., Sawitzki G., Smith C., Smyth G., Tierney L., Yang J.Y. & Zhang J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome biology* 5, R80.

Georges M. (2007) Mapping, fine mapping, and molecular dissection of quantitative trait Loci in domestic animals. *Annual review of genomics and human genetics* 8, 131-162.

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. *Journal of animal science* 82, 17-31.

Goddard M.E. & Hayes B.J. (2007) Genomic selection. *Journal of Animal Breeding and Genetics* 124, 323-330.

Goliasova E. & Wolf J. (2004) Impact of the ESR gene on litter size and production traits in Czech Large White pigs. *Animal Genetics* 35, 293-297.

Gonzalez-Ramon N., Hoebe K., Alava M.A., Van Leengoed L., Pineiro M., Carmona S., Iturralde M., Lampreave F. & Pineiro A. (2000) Pig MAP/ITIH4 and haptoglobin are interleukin-6-dependent acute-phase plasma proteins in porcine primary cultured hepatocytes. *European journal of biochemistry / FEBS* 267, 1878-1885.

Govindarajan B. & Gipson I.K. (2010) Membrane-tethered mucins have multiple functions on the ocular surface. *Experimental eye research* 90, 655-663.

Grimson A., Srivastava M., Fahey B., Woodcroft B.J., Chiang H.R., King N., Degnan B.M., Rokhsar D.S. & Bartel D.P. (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* 455, 1193-1197.

Groenen, M., Lawrence B., Schook & Archibald A. (2011). Pig Genomics. In: M. F. Rothschild and A. Ruvinsky (eds.) The genetics of the pig. p 179-200.

Halbeisen R.E., Galgano A., Scherrer T. & Gerber A.P. (2008) Post-transcriptional gene regulation: from genome-wide studies to principles. *Cellular and molecular life sciences : CMLS* 65, 798-813.

Haley , C.S., Alvalos, E. and Smith, C. (1988) Selection for litter size in the pig. *Animal Breeding Abstract* 56, 317-332.

Hawkins S.M., Buchold G.M. & Matzuk M.M. (2011) Minireview: The Roles of Small RNA Pathways in Reproductive Medicine. *Molecular endocrinology (Baltimore, Md.)*.

Heimberg A.M., Cowper-Sal-lari R., Semon M., Donoghue P.C. & Peterson K.J. (2010) microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proceedings of the National Academy of Sciences of the United States of America* 107, 19379-19383.

Heimberg A.M., Sempere L.F., Moy V.N., Donoghue P.C. & Peterson K.J. (2008) MicroRNAs and the advent of vertebrate morphological complexity. *Proceedings of the National Academy of Sciences of the United States of America* 105, 2946-2950.

Hermesch S. (2000) Genetic parameters for lean meat yield, meat quality, reproduction and feed efficiency traits for Australian pigs:: 2. Genetic relationships between production, carcase and meat quality traits. *Livestock Production Science* 65, 249.

Herpin P., Hulin J.C., Le Dividich J. & Fillaut M. (2001) Effect of oxygen inhalation at birth on the reduction of early postnatal mortality in pigs. *Journal of animal science* 79, 5-10.

Herpin P., Damon M. & Le Dividich J. (2002) Development of thermoregulation and neonatal survival in pigs. *Livestock Production Science* 78, 25-45.

Hettinger A.M., Allen M.R., Zhang B.R., Goad D.W., Malayer J.R. & Geisert R.D. (2001) Presence of the acute phase protein, bikunin, in the endometrium of gilts during estrous cycle and early pregnancy. *Biology of reproduction* 65, 507-513.

Himmelfarb M., Klopocki E., Grube S., Staub E., Klaman I., Hinzmann B., Kristiansen G., Rosenthal A., Durst M. & Dahl E. (2004) ITIH5, a novel member of the inter-alpha-trypsin inhibitor heavy chain family is downregulated in breast cancer. *Cancer letters* 204, 69-77.

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. *Journal of animal science* 82, 3421-3429.

Holl J.W. & Robison O.W. (2003) Results from nine generations of selection for increased litter size in swine. *Journal of animal science* 81, 624-629.

Holm B., Bakken M., Klemetsdal G. & Vangen O. (2004) Genetic correlations between reproduction and production traits in swine. *Journal of animal science* 82, 3458-3464.

Hong X., Luense L.J., McGinnis L.K., Nothnick W.B. & Christenson L.K. (2008) Dicer1 is essential for female fertility and normal development of the female reproductive system. *Endocrinology* 149, 6207-6212.

Horak P., Urban T. & Dvorak J. (2005) The FUT1 and ESR genes--their variability and associations with reproduction in Prestice Black-Pied sows. *Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie* 122, 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. *Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie* 122, 56-61.

Huby, M., Gogué, J.M., Maignel, L. and Bidanel, J.P. (2003) Genetic correlations between litter size and weights, piglet weight variability and piglet survival from birth to weaning in Large White pigs. Journées de la Recherche Porcine en France 35, 293–300.

Hunter M.G., Robinson R.S., Mann G.E. & Webb R. (2004) Endocrine and paracrine control of follicular development and ovulation rate in farm species. *Animal Reproduction Science* 82-83, 461-477.

Huntzinger E. & Izaurralde E. (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature reviews.Genetics* 12, 99-110.

Idris N. & Carraway K.L. (2000) Regulation of sialomucin complex/Muc4 expression in rat uterine luminal epithelial cells by transforming growth factor-beta: implications for blastocyst implantation. *Journal of cellular physiology* 185, 310-316.

Idris N. & Carraway K.L. (1999) Sialomucin complex (Muc4) expression in the rat female reproductive tract. *Biology of reproduction* 61, 1431-1438.

Irgang R., Favero J.A. & Kennedy B.W. (1994) Genetic parameters for litter size of different parities in Duroc, Landrace, and large white sows. *Journal of animal science* 72, 2237-2246.

Ishida J., Matsuoka T., Saito-Fujita T., Inaba S., Kunita S., Sugiyama F., Yagami K. & Fukamizu A. (2011) Pregnancy-associated homeostasis and dysregulation: lessons from genetically modified animal models. *Journal of Biochemistry* 150, 5-14.

Jacobsen M., Kracht S.S., Esteso G., Cirera S., Edfors I., Archibald A.L., Bendixen C., Andersson L., Fredholm M. & Jørgensen C.B. (2010) Refined candidate region specified by haplotype sharing for Escherichia coli F4ab/F4ac susceptibility alleles in pigs. *Animal Genetics* 41, 21-25.

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 / National Research Council Canada = Genome / Conseil national de recherches Canada* 44, 7-12.

Johnson R.K., Zimmerman D.R. & Kittok R.J. (1984) Selection for components of reproduction in swine. *Livestock Production Science* 11, 541-558.

Jones, G. F. (1998). Genetic aspects of domestication, common breeds, and their origin. In: M. F. Rothschild and A. Ruvinsky (eds.) The genetics of the pig. p 17-50.

Jorgensen C.B., Cirera S., Anderson S.I., Archibald A.L., Raudsepp T., Chowdhary B., Edfors-Lilja I., Andersson L. & Fredholm M. (2003) Linkage and comparative mapping of the locus controlling susceptibility towards E. COLI F4ab/ac diarrhoea in pigs. *Cytogenetic and genome research* 102, 157-162.

Ka H., Seo H., Kim M., Moon S., Kim H. & Lee C.K. (2008) Gene expression profiling of the uterus with embryos cloned by somatic cell nuclear transfer on day 30 of pregnancy. *Animal Reproduction Science* 108, 79-91.

Kanis E., De Greef K.H., Hiemstra A. & van Arendonk J.A. (2005) Breeding for societally important traits in pigs. *Journal of animal science* 83, 948-957.

Kaufmann D., Hofer A., Bidanel J.P. & Künzi N. (2000) Genetic parameters for individual birth and weaning weight and for litter size of Large White pigs; Genetische Parameter für Geburtsgewicht, Absetzgewicht und Würfgrösse bei Large White Schweinen. *Journal of Animal Breeding and Genetics* 117, 121-128.

Keene J.D. (2007) RNA regulons: coordination of post-transcriptional events. *Nature reviews.Genetics* 8, 533-543.

Keys J.L. & King G.J. (1990) Microscopic examination of porcine conceptus-maternal interface between days 10 and 19 of pregnancy. *The American Journal of Anatomy* 188, 221-238.

Kim D., Song J., Kim S., Kang S.S. & Jin E.J. (2011) MicroRNA-142-3p regulates TGF-beta3-mediated region-dependent chondrogenesis by regulating ADAM9. *Biochemical and biophysical research communications* 414, 653-659.

Kim S.Y., Romero R., Tarca A.L., Bhatti G., Lee J., Chaiworapongsa T., Hassan S.S. & Kim C.J. miR-143 Regulation of Prostaglandin-Endoperoxidase Synthase 2 in the Amnion: Implications for Human Parturition at Term. - *PLoS ONE*, - e24131.

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. *Biology of reproduction* 68, 2172-2179.

Kirkpatrick, B.W., (2002). QTL and candidate gene effects on reproduction in livestock: Progress and prospects. In: Proceedings of the Seventh World Congress on Genetics Applied to Livestock Production, Montpellier, vol. 30, pp. 633–636.

Knol E.F., Leenhouwers J.I. & van der Lende T. (2002) Genetic aspects of piglet survival. *Livestock Production Science* 78, 47-55.

Komatsu M., Carraway C.A., Fregien N.L. & Carraway K.L. (1997) Reversible disruption of cell-matrix and cell-cell interactions by overexpression of sialomucin complex. *The Journal of biological chemistry* 272, 33245-33254.

Korwin-Kossakowska, A., Kamyczek, M., Cieslak, D., Pierzchaa, M. and Kury, J. (2002) The effect of leptin (LEP), leptin receptor (LEPR) and osteopontin (OPN) gene polymorphism on selected reproduction traits of synthetic Line 990 sows. Animal Science Papers and Reports 20, 159–168.

Koscinski I., Viville S., Porchet N., Bernigaud A., Escande F., Defossez A. & Buisine M.P. (2006) MUC4 gene polymorphism and expression in women with implantation failure. *Human reproduction (Oxford, England)* 21, 2238-2245.

Krol J., Loedige I. & Filipowicz W. (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews.Genetics* 11, 597-610.

Kumar P., Johnston B.H. & Kazakov S.A. (2011) miR-ID: a novel, circularization-based platform for detection of microRNAs. *RNA (New York, N.Y.)* 17, 365-380.

Lagow E., DeSouza M.M. & Carson D.D. (1999) Mammalian reproductive tract mucins. *Human reproduction update* 5, 280-292.

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. *Journal of animal science* 69, 3129-3143.

Lander E.S. & Botstein D. (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. *Genetics* 121, 185-199.

Larson G., Cucchi T., Fujita M., Matisoo-Smith E., Robins J., Anderson A., Rolett B., Spriggs M., Dolman G., Kim T.H., Thuy N.T., Randi E., Doherty M., Due R.A., Bollt R., Djubiantono T., Griffin B., Intoh M., Keane E., Kirch P., Li K.T., Morwood M., Pedrina L.M., Piper P.J., Rabett R.J., Shooter P., Van den Bergh G., West E., Wickler S., Yuan J., Cooper A. & Dobney K. (2007) Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4834-4839.

Le Dividich J., Mormede P., Catheline M. & Caritez J.C. (1991) Body composition and cold resistance of the neonatal pig from European (Large White) and Chinese (Meishan) breeds. *Biology of the neonate* 59, 268-277.

Lee R.C., Feinbaum R.L. & Ambros V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843-854.

Lei L., Jin S., Gonzalez G., Behringer R.R. & Woodruff T.K. (2010) The regulatory role of Dicer in folliculogenesis in mice. *Molecular and cellular endocrinology* 315, 63-73.

Lei B., Gao S., Luo L.F., Xia X.Y., Jiang S.W., Deng C.Y., Xiong Y.Z. & Li F.E. (2011) A SNP in the miR-27a gene is associated with litter size in pigs. *Molecular biology reports* 38, 3725-3729.

Levine M. & Tjian R. (2003) Transcription regulation and animal diversity. *Nature* 424, 147-151.

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 ıChinese Erhualian resource population. *Animal Genetics* 40, 963-966.

Li N., Y.F. Zhao, L. Xiao, F.J. Zhang, Y.Z. Chen (1998) Candidate gene approach for identification of genetic loci controlling litter size in pigs Proceedings of the Sixth World Congress Genet. Livest. Prod., vol. 26, Armidale, Australia, 11–16 January, pp. 183–190

Li M., Liu Y., Wang T., Guan J., Luo Z., Chen H., Wang X., Chen L., Ma J., Mu Z., Jiang A.A., Zhu L., Lang Q., Zhou X., Wang J., Zeng W., Li N., Li K., Gao X. & Li X. (2011) Repertoire of porcine microRNAs in adult ovary and testis by deep sequencing. *International journal of biological sciences* 7, 1045-1055.

Li, F.E., Mei, S.Q., Deng, C.Y., Jiang, S.W., Zuo, B., Zheng, R., Li, J.L., Xu, D.Q., Lei, M.G. and Xiong, Y.Z. (2008) Association of a microsatellite flanking FSHB gene with reproductive traits and reproductive tract components in pigs. Czech Journal of Animal Science 53, 139–144.

Linsen S.E., de Wit E., Janssens G., Heater S., Chapman L., Parkin R.K., Fritz B., Wyman S.K., de Bruijn E., Voest E.E., Kuersten S., Tewari M. & Cuppen E. (2009) Limitations and possibilities of small RNA digital gene expression profiling. *Nature methods* 6, 474-476.

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)* 25, 402-408.

Marques J.T., Kim K., Wu P.H., Alleyne T.M., Jafari N. & Carthew R.W. (2010) Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nature structural & molecular biology* 17, 24-30.

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. *Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie* 128, 344-353.

Mata J., Marguerat S. & Bahler J. (2005) Post-transcriptional control of gene expression: a genome-wide perspective. *Trends in biochemical sciences* 30, 506-514.

Matsuzaki S. (2011) DNA microarray analysis in endometriosis for development of more effective targeted therapies. *Frontiers in bioscience (Elite edition)* 3, 1139-1153.

McNeer R.R., Carraway C.A., Fregien N.L. & Carraway K.L. (1998a) Characterization of the expression and steroid hormone control of sialomucin complex in the rat uterus: implications for uterine receptivity. *Journal of cellular physiology* 176, 110-119.

McNeer R.R., Huang D., Fregien N.L. & Carraway K.L. (1998b) Sialomucin complex in the rat respiratory tract: a model for its role in epithelial protection. *The Biochemical journal* 330 (Pt 2), 737-744.

Merks, J.W.M. (2000) One century of genetic changes in pigs and the future needs. In: Hill, W.G., Bishop, S.C., McGuirk, B., McKay, J.C., Simm, G. and Webb, A.J. (eds) The Challenge of Genetic Change in Animal Production. British Society of Animal Science, Edinburgh, pp. 8–19.

Mestdagh P., Feys T., Bernard N., Guenther S., Chen C., Speleman F. & Vandesompele J. (2008) High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. *Nucleic acids research* 36, e143.

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

Milligan B.N., Fraser D. & Kramer D.L. (2002) Within-litter birth weight variation in the domestic pig and its relation to pre-weaning survival, weight gain, and variation in weaning weights. *Livestock Production Science* 76, 181-191.

Moav, R. 1966. Specialized sire and dam lines. Animal Production 8: 193, 203 and 365

Moniaux N., Escande F., Porchet N., Aubert J.P. & Batra S.K. (2001) Structural organization and classification of the human mucin genes. *Frontiers in bioscience : a journal and virtual library* 6, D1192-206.

Moniaux N., Nollet S., Porchet N., Degand P., Laine A. & Aubert J.P. (1999) Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. *The Biochemical journal* 338 (Pt 2), 325-333.

Montenegro D., Romero R., Kim S.S., Tarca A.L., Draghici S., Kusanovic J.P., Kim J.S., Lee D.C., Erez O., Gotsch F., Hassan S.S. & Kim C.J. (2009) Expression patterns of microRNAs in the chorioamniotic membranes: a role for microRNAs in human pregnancy and parturition. *The Journal of pathology* 217, 113-121.

Morris A.R., Mukherjee N. & Keene J.D. (2010) Systematic analysis of posttranscriptional gene expression. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 2, 162-180.

Mouillet J.F., Chu T. & Sadovsky Y. (2011) Expression patterns of placental microRNAs. *Birth defects research.Part A, Clinical and molecular teratology* 91, 737-743.

Munoz G., Ovilo C., Silio L., Tomas A., Noguera J.L. & Rodriguez M.C. (2009) Single- and joint-population analyses of two experimental pig crosses to confirm quantitative trait loci on Sus scrofa chromosome 6 and leptin receptor effects on fatness and growth traits. *Journal of animal science* 87, 459-468.

Munoz G., Ovilo C., Estelle J., Silio L., Fernandez A. & Rodriguez C. (2007) Association with litter size of new polymorphisms on ESR1 and ESR2 genes in a Chinese-European pig line. *Genetics, selection, evolution : GSE* 39, 195-206.

Muñoz M., (2007) Estudio de los efectos del gen RBP4 sobre el tamaño camada en una línea sintética de origen Chino-Europeo y en un cruce experimental Ibérico x Meishan. Trabajo de investigación para la obtención de la suficiencia investigadora (DEA).

Murchison E.P., Stein P., Xuan Z., Pan H., Zhang M.Q., Schultz R.M. & Hannon G.J. (2007) Critical roles for Dicer in the female germline. *Genes & development* 21, 682-693.

Nagaraja A.K., Andreu-Vieyra C., Franco H.L., Ma L., Chen R., Han D.Y., Zhu H., Agno J.E., Gunaratne P.H., DeMayo F.J. & Matzuk M.M. (2008) Deletion of Dicer in somatic cells of the female reproductive tract causes sterility. *Molecular endocrinology (Baltimore, Md.)* 22, 2336-2352.

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. *Animal Reproduction Science* 110, 283-292.

Noguera J.L., Rodriguez C., Varona L., Tomas A., Munoz 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., 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. *Journal of animal science* 80, 2548-2555.

Nonneman D.J., Wise T.H., Ford J.J., Kuehn L.A. & Rohrer G.A. (2006) Characterization of the aldo-keto reductase 1C gene cluster on pig chromosome 10: possible associations with reproductive traits. *BMC veterinary research* 2, 28.

Nothnick W. (2012) The role of microRNAs in the female reproductive tract. *Reproduction (Cambridge, England)*.

Ohlsson Teague E.M., Van der Hoek K.H., Van der Hoek M.B., Perry N., Wagaarachchi P., Robertson S.A., Print C.G. & Hull L.M. (2009) MicroRNA-regulated pathways associated with endometriosis. *Molecular endocrinology (Baltimore, Md.)* 23, 265-275.

Okamura K., Phillips M.D., Tyler D.M., Duan H., Chou Y.T. & Lai E.C. (2008) The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nature structural & molecular biology* 15, 354-363.

Orphanides G. & Reinberg D. (2002) A Unified Theory of Gene Expression. *Cell* 108, 439-451.

Ostrup E., Bauersachs S., Blum H., Wolf E. & Hyttel P. (2010) Differential endometrial gene expression in pregnant and nonpregnant sows. *Biology of reproduction* 83, 277-285.

Otsuka M., Zheng M., Hayashi M., Lee J.D., Yoshino O., Lin S. & Han J. (2008) Impaired microRNA processing causes corpus luteum insufficiency and infertility in mice. *The Journal of clinical investigation* 118, 1944-1954.

Pan Q. & Chegini N. (2008) MicroRNA signature and regulatory functions in the endometrium during normal and disease states. *Seminars in reproductive medicine* 26, 479-493.

Pan Q., Luo X., Toloubeydokhti T. & Chegini N. (2007) The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. *Molecular human reproduction* 13, 797-806.

Pasquinelli A.E., Reinhart B.J., Slack F., Martindale M.Q., Kuroda M.I., Maller B., Hayward D.C., Ball E.E., Degnan B., Muller P., Spring J., Srinivasan A., Fishman M., Finnerty J., Corbo J., Levine M., Leahy P., Davidson E. & Ruvkun G. (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86-89.

Peng Q.L., Ren J., Yan X.M., Huang X., Tang H., Wang Y.Z., Zhang B. & Huang L.S. (2007) The g.243A>G mutation in intron 17 of MUC4 is significantly associated with susceptibility/resistance to ETEC F4ab/ac infection in pigs. *Animal Genetics* 38, 397-400.

Perez-Enciso M. & Misztal I. (2004) Qxpak: a versatile mixed model application for genetical genomics and QTL analyses. *Bioinformatics (Oxford, England)* 20, 2792-2798.

Perez-Enciso M. & Gianola D. (1992) Estimates of genetic parameters for litter size in six strains of Iberian pigs. *Livestock Production Science* 32, 283-293.

Peterson K.J., Dietrich M.R. & McPeek M.A. (2009) MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. *BioEssays : news and reviews in molecular, cellular and developmental biology* 31, 736-747.

Pineiro M., Alava M.A., Gonzalez-Ramon N., Osada J., Lasierra P., Larrad L., Pineiro A. & Lampreave F. (1999) ITIH4 serum concentration increases during acute-phase processes in human patients and is up-regulated by interleukin-6 in hepatocarcinoma HepG2 cells. *Biochemical and biophysical research communications* 263, 224-229.

Pope, W. F. 1994. Embryonic mortality in swine. In: R. D. Geisert and M. T. Zavy (ed.) Embryonic Mortality in Domestic Species.

Ramirez O., Tomas A., Barragan C., Noguera J.L., Amills M. & Varona L. (2009) Pig melatonin receptor 1a (MTNR1A) genotype is associated with seasonal variation of sow litter size. *Animal Reproduction Science* 115, 317-322.

Ramirez O., Tomàs A., Casellas J., Blanch M., Noguera J. & Amills M. (2008) An Association Analysis Between a Silent C558T Polymorphism at the Pig Vascular Cell Adhesion Molecule 1 Locus and Sow Reproduction and Piglet Survivability Traits. *Reproduction in Domestic Animals* 43, 542-546.

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, e6524.

Rampoldi A., Jacobsen M.J., Bertschinger H.U., Joller D., Burgi E., Vogeli P., Andersson L., Archibald A.L., Fredholm M., Jorgensen C.B. & Neuenschwander S. (2011) The receptor locus for Escherichia coli F4ab/F4ac in the pig maps distal to the MUC4-LMLN region. *Mammalian genome : official journal of the International Mammalian Genome Society* 22, 122-129.

Ro S., Park C., Young D., Sanders K.M. & Yan W. (2007) Tissue-dependent paired expression of miRNAs. *Nucleic acids research* 35, 5944-5953.

Rodriguez C., Tomas A., Alves E., Ramirez O., Arque M., Munoz 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. *Animal Genetics* 36, 490-496.

Roehe, R. & Kalm, E. (2000) Estimation of genetic and environmental risk factors associated with pre-weaning mortality in piglets using generalized linear mixed models. Animal Science 70, 227–240.

Roehe R. & Kennedy B.W. (1993) Effect of selection for maternal and direct genetic effects on genetic improvement of litter size in swine. *Journal of animal science* 71, 2891-2904.

Ron M. & Weller J.I. (2007) From QTL to QTN identification in livestock ? winning by points rather than knock-out: a review. *Animal Genetics* 38, 429-439.

Ronaghi M., Pettersson B., Uhlen M. & Nyren P. (1998) PCR-introduced loop structure as primer in DNA sequencing. *BioTechniques* 25, 876-8, 880-2, 884.

Rosendo A., Canario L., Druet T., Gogue J. & Bidanel J.P. (2007a) Correlated responses of pre- and postweaning growth and backfat thickness to six generations of selection for ovulation rate or prenatal survival in French Large White pigs. *Journal of animal science* 85, 3209-3217.

Rosendo A., Druet T., Gogué J. & Bidanel J.P. (2007b) Direct responses to six generations of selection for ovulation rate or prenatal survival in Large White pigs. *Journal of animal science* 85, 356-364.

Ross J.W., Ashworth M.D., Stein D.R., Couture O.P., Tuggle C.K. & Geisert R.D. (2009) Identification of differential gene expression during porcine conceptus rapid trophoblastic elongation and attachment to uterine luminal epithelium. *Physiological genomics* 36, 140-148.

Rothschild M., Jacobson C., Vaske D., Tuggle C., Wang L., Short T., Eckardt G., Sasaki S., Vincent A., McLaren D., Southwood O., van der Steen H., Mileham A. & Plastow G. (1996) The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proceedings of the National Academy of Sciences* 93, 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. *Mammalian genome : official journal of the International Mammalian Genome Society* 11, 75-77.

CAB International, Oxon.Ruan Y., Le Ber P., Ng H.H. & Liu E.T. (2004) Interrogating the transcriptome. *Trends in biotechnology* 22, 23-30.

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. *Journal of animal science* 79, 2286-2297.

Salier J.P., Rouet P., Raguenez G. & Daveau M. (1996) The inter-alpha-inhibitor family: from structure to regulation. *The Biochemical journal* 315 (Pt 1), 1-9.

Salustri A., Yanagishita M. & Hascall V.C. (1989) Synthesis and accumulation of hyaluronic acid and proteoglycans in the mouse cumulus cell-oocyte complex during follicle-stimulating hormone-induced mucification. *The Journal of biological chemistry* 264, 13840-13847.

Scarchilli L., Camaioni A., Bottazzi B., Negri V., Doni A., Deban L., Bastone A., Salvatori G., Mantovani A., Siracusa G. & Salustri A. (2007) PTX3 interacts with inter-alpha-trypsin inhibitor: implications for hyaluronan organization and cumulus oophorus expansion. *The Journal of biological chemistry* 282, 30161-30170.

See M.T., Mabry J.W. & Bertrand J.K. (1993) Restricted maximum likelihood estimation of variance components from field data for number of pigs born alive. *Journal of animal science* 71, 2905-2909.

Shi R. & Chiang V.L. (2005) Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques* 39, 519-525.

Short T.H., Rothschild M.F., Southwood O.I., McLaren D.G., de Vries A., van der Steen 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. *Journal of animal science* 75, 3138-3142.

Silió, L., Rodrigáñez, J. & Toro, M.A., (2001). La selección de cerdos ibéricos. In: Buxadé, C., Daza, A. (Eds.), Porcino Ibérico: Aspectos Claves. Madrid, pp. 125–149.

Silva P.V., Guimaraes S.E., Guimaraes J.D., Neto J.B., Lopes P.S., do Nascimento C.S., de Campos C.F., Weller M.M., Botelho M.E. & Faria V.R. (2011) Gene expression in swine granulosa cells and ovarian tissue during the estrous cycle. *Genetics and molecular research : GMR* 10, 2258-2267.

Sirotkin A.V., Ovcharenko D., Grossmann R., Lauková M. & Mlyn?ek M. (2009) Identification of MicroRNAs controlling human ovarian cell steroidogenesis via a genome-scale screen. *Journal of cellular physiology* 219, 415-420.

Smith, C. 1964. The use of specialised sire and dam lines in selection for meat production. Animal Production 6 : 337

Solanes F.X., Grandinson K., Rydhmer L., Stern S., Andersson K. & Lundeheim N. (2004) Direct and maternal influences on the early growth, fattening performance, and carcass traits of pigs. *Livestock Production Science* 88, 199-212.

Southwood O.I. & Kennedy B.W. (1990) Estimation of direct and maternal genetic variance for litter size in Canadian Yorkshire and Landrace swine using an animal model. *Journal of animal science* 68, 1841-1847.

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. *Reproduction in domestic animals = Zuchthygiene* 44, 100-105.

Spotter A. & Distl O. (2006) Genetic approaches to the improvement of fertility traits in the pig. *Veterinary journal (London, England : 1997)* 172, 234-247.

Srivastava S.K., Bhardwaj A., Singh S., Arora S., Wang B., Grizzle W.E. & Singh A.P. (2011) MicroRNA-150 directly targets MUC4 and suppresses growth and malignant behavior of pancreatic cancer cells. *Carcinogenesis* 32, 1832-1839.

Strachan T. & Read A. (2011). Human molecular Genetics

Su L., Zhao S., Zhu M. & Yu M. (2010) Differential expression of microRNAs in porcine placentas on Days 30 and 90 of gestation. *Reproduction, Fertility and Development* 22, 1175-1182.

Tabor H.K., Risch N.J. & Myers R.M. (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature reviews.Genetics* 3, 391-397.

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. *The Journal of reproduction and development* 53, 455-463.

Tomas A., Casellas J., Ramirez O., Munoz G., Noguera J.L. & Sanchez A. (2006) High amino acid variation in the intracellular domain of the pig prolactin receptor (PRLR) and its relation to ovulation rate and piglet survival traits. *Journal of animal science* 84, 1991-1998.

Tomás A., Ramírez O., Casellas J., Muñoz G., Sánchez A., Barragán C., Arqué M., Riart I., Óvilo C., Noguera J.L., Amills M. & Rodríguez C. (2011) Quantitative trait loci for fatness at growing and reproductive stages in Iberian ıMeishan F2 sows. *Animal Genetics* 42, 548-551.

Tomás A., Casellas J., Ramírez O., Pérez-Enciso M., Rodríguez C., Noguera J.L. & Sánchez A. (2006a) Polymorphisms of the porcine dopamine?-hydroxylase gene and their relation to reproduction and piglet survivability in an Iberian ıMeishan F2 intercross. *Animal Genetics* 37, 279-282.

Tomás A., Frigo E., Casellas J., Ramírez O., Óvilo C., Noguera J.L. & Sánchez A. (2006b) An association study between polymorphisms of the porcine bone morphogenetic protein receptor type1?(BMPR1B) and reproductive performance of Iberian ıMeishan F2 sows. *Animal Genetics* 37, 297-298.

Tribout T., Iannuccelli N., Druet T., Gilbert H., Riquet J., Gueblez R., Mercat M.J., Bidanel J.P., Milan D. & Le Roy P. (2008) Detection of quantitative trait loci for reproduction and production traits in Large White and French Landrace pig populations. *Genetics, selection, evolution : GSE* 40, 61-78.

Tribout, T. and Bidanel, J.P. (2008a) Genetic parameters of number of piglets born alive and relationships with on-farm performance traits in French Landrace and Large White pig breeds. *Journées de la Recherche Porcine* 40, 113–118.

Tribout, T., Caritez, J.C., Gogué, J., Gruand, J., Billon, Y., Le Dividich, J., Quesnel, H. and Bidanel, J.P. (2003) Estimation of realised genetic trend in French Large White pigs from 1977 to 1998 using frozen semen: results for some female reproduction traits. *Journées de la Recherche Porcine en France* 35, 285–292.

Vallet J.L., Freking B.A., Leymaster K.A. & Christenson R.K. (2005) Allelic variation in the erythropoietin receptor gene is associated with uterine capacity and litter size in swine. *Animal Genetics* 36, 97-103.

Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M, Andersson L. (2003)

A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. Nature. 425, 832-836.

van Rens B.T., Evans G.J. & van der Lende T. (2003) Components of litter size in gilts with different prolactin receptor genotypes. *Theriogenology* 59, 915-926.

van Rens B.T., 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, 1635-1649.

van Rens B.T. & van der Lende T. (2002) Litter size and piglet traits of gilts with different prolactin receptor genotypes. *Theriogenology* 57, 883-893.

Varona, L. & Noguera, J.L. (2001) Variance components of fertility in Spanish Landrace pigs. *Livestock Production Science* 67, 217–221.

Vasudevan S. (2011) Posttranscriptional Upregulation by MicroRNAs. *Wiley Interdisciplinary Reviews: RNA*, n/a-n/a.

Vasudevan S., Tong Y. & Steitz J.A. (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science (New York, N.Y.)* 318, 1931-1934.

A.L. Vincent, G. Evans, T.H. Short, O.I. Southwood, G.S. Plastow, C.K. Tuggle, M.F. Rothschild. (1998) The prolactin receptor gene is associated with increased litter size in pigs Proceedings of the Sixth World Congress Genet. Livest. Prod., vol. 27, Armidale, Australia, 11–16 January, pp. 15–18

Wheeler B.M., Heimberg A.M., Moy V.N., Sperling E.A., Holstein T.W., Heber S. & Peterson K.J. (2009) The deep evolution of metazoan microRNAs. *Evolution & development* 11, 50-68.

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. *Mammalian genome : official journal of the International Mammalian Genome Society* 10, 573-578.

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. *Journal of animal science* 77, 1654-1658.

Wu C.I., Shen Y. & Tang T. (2009) Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome research* 19, 734-743.

Yao N., Yang B.Q., Liu Y., Tan X.Y., Lu C.L., Yuan X.H. & Ma X. (2010) Follicle-stimulating hormone regulation of microRNA expression on progesterone production in cultured rat granulosa cells. *Endocrine* 38, 158-166.

Zhang S., Bidanel J.P., Burlot T., Legault C. & Naveau J. (2000) Genetic parameters and genetic trends in the Chinese x European Tiameslan composite pig line. I. Genetic parameters. *Genetics selection evolution* 32, 41-56.

Zhao H., Rothschild M.F., Fernando R.L. & Dekkers J.C. (2003) Tests of candidate genes in breed cross populations for QTL mapping in livestock. *Mammalian genome : official journal of the International Mammalian Genome Society* 14, 472-482.

Zhou Y, Zhu YZ, Zhang SH, Wang HM, Wang SY, Yang XK (2011) MicroRNA expression profiles in premature ovarian failure patients and its potential regulate functions. Chinese journal of birth health and heredity 19:20-22.

Zhu M. & Zhao S. (2007) Candidate gene identification approach: progress and challenges. *International journal of biological sciences* 3, 420-427.

Zhuo L. & Kimata K. (2008) Structure and function of inter-alpha-trypsin inhibitor heavy chains. *Connective tissue research* 49, 311-320.