UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Ciència Animal i dels Aliments

Facultat de Veterinària

CENTRE DE RECERCA EN AGRIGENÒMICA

Departament de Genètica Animal

"Genetic and molecular basis of reproductive efficiency in swine"

Sarai Córdoba Terreros

PhD Thesis in Animal Production

Bellaterra, 2015

Supervisor:

Dr. Armand Sánchez Bonastre

Universitat Autònoma de Barcelona

“It is a miracle that curiosity survives formal education.”

- Albert Einstein (1879-1955)

A **mis padres**, por su incondicionalidad.

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

CERTIFICA:

Que la **Sarai Córdoba Terreros** ha realitzat sota la seva direcció el treball de recerca “Genetic and molecular basis of reproductive efficiency in swine” 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 i a la unitat de Genètica Animal del Centre de Recerca en Agrigenòmica.

Bellaterra, a 9 d’Octubre de 2015

Dr. Armand Sánchez Bonastre

SUMMARY/RESUM 13

List of Tables 17

List of Figures 19

List of publications 21

Related publications by the author 22

Abbreviations 23

1. GENERAL INTRODUCTION 29

1.1. Pig production 29

1.1.1. Improvement of pig production 30

1.2. Reproduction in swine 30

1.2.1. Reproductive cycle of a sow 30

1.2.1.1. Follicular phase 31

1.2.1.2. Luteal phase 31

1.2.1.3. Pregnancy establishment 32

1.3. Genetic basis of swine reproduction 33

1.3.1. Litter size 34

1.3.2. Reproductive QTLs 35

1.3.3. Candidate genes for litter size in pigs 37

1.4. Pig genomics 38

1.4.1. Transcriptome profiling 39

1.4.1.1. High-throughput sequencing (RNA-sequencing) 39

1.4.2. Gene expression estimation 46

1.4.2.1. Gene expression quantification by RT-qPCR 47

1.4.2.2. MicroRNA expression quantification by RT-qPCR 49

1.5. Regulation of gene expression 51

1.5.1. MicroRNAs 52

1.5.1.1. MiRNA targeting 53

1.5.1.2. Functional validation of miRNA targeting 55

1.5.2. Regulation of miRNAs processing 56

1.5.3. Role of miRNAs in reproduction 57

1.6. MEIBMAP intercross 57

1.6.1. QTLs identified within the MEIBMAP project 59

1.6.2. Candidate genes identified within the MEIBMAP project 60

2. OBJECTIVES 65

3. ARTICLES AND STUDIES 69

3.1. Endometrial transcriptome profiling studies

3.1.1. **Study I**: Analysis of gene expression differences between extreme prolificacy phenotypes 69

3.1.2. **Study II**: Validation of reproduction-related candidate genes 87

3.2. Gene expression regulation studies

3.2.1. **Study III**: Determination of polymorphisms affecting the regulatory function of reproductive-related miRNAs 101

3.2.2. **Study IV**: Functional validation of the miRNA-mRNA interaction 123

4. GENERAL DISCUSSION 143

5. CONCLUSIONS 159

6. REFERENCES 163

7. ANNEXES 189

8. AKNOWLEDGMENTS 221

In recent years, transcriptome characterization has seen a remarkable rise, becoming a hot topic in genomic research either in human or animal genetics. In this last, advances in transcriptomics have addressed the goal to better understand those traits with higher economic impact. One of the most important species in livestock production are pigs. Reproductive traits such as prolificacy can directly impact porcine profitability, but large genetic variation and low heritability have been found regarding litter size among porcine breeds. This highlights the importance to perform expression profiling experiments in porcine breeds with extreme prolificacy phenotypes, to better understand those gene interactions and regulatory mechanisms affecting litter size in pigs.

In this thesis, we provide a global view of the endometrial transcriptome of two porcine breeds that differ significantly in their prolificacy levels, giving a list of more than one hundred differentially expressed genes associated with critical steps of embryonic survival during sow's gestation. These expression differences have been validated for 12 genes providing a list of new candidate genes that may play key role on the genetic architecture of prolificacy-related traits in pigs. We hypothesized that the observed differences in the expression level of these genes between Iberian x Meishan F_2 sows with divergent prolificacy phenotypes might respond to a different expression pattern of microRNAs (miRNAs), known to function as post-transcriptional down-regulators of gene expression. To validate this hypothesis, we explored the endometrial miRNA expression profile by RNA-seq identifying 10 differentially expressed miRNAs. Expression levels appear to be similar after relative quantification, despite significant correlations were found between the expression of *ssc-miR-92a* and *ssc-miR-133a* and candidate genes *MMP8*, *PTGS2*, *PTHLH* and *SCNN1G*. We functionally characterized nine reproduction-related miRNAs identifying a total of 13 SNPs in their precursor sequences. To determine the effect of these variants in the reproductive efficiency of the pregnant sows, we performed an association study that revealed that the genotype for the variants in *ssc-mir-27a*, *ssc-mir-29b-2* and *ssc-mir-106a* was determinant for the mature miRNA expression levels and the EBVs. Finally, a functional validation of the miRNA-mediated regulation of *ADM*, *HTRA3, PTHLH* and *VEGFA* upon they target miRNAs *ssc-miR-181d-5p*, *ssc-miR-101-3p*, *ssc-miR-144* and *ssc-miR-195-5p* respectively, allowed us to find a direct relationship between these interactions and decreased levels of gene expression.

En els darrers anys, la caracterització del transcriptoma s'ha convertit en un tema candent a la recerca genòmica, ja sigui en humans o en animals. En aquests últims, els avanços en transcriptòmica tenen com a principal objectiu entendre millor els caràcters amb major impacte econòmic. Una de les espècies més importants en la producció ramadera és la porcina. Els caràcters reproductius com la prolificitat poden afectar directament la seva rendibilitat, però la gran variabilitat genètica existent entre races porcines i la baixa heretabilitat d'aquest caràcter han fet de la seva selecció tot un repte. Això posa de manifest la importància d'estudiar les interaccions gèniques i els mecanismes de regulació que afecten el tamany final de la camada en aquesta espècie.

En aquesta tesi, oferim una visió global del transcriptoma de l'endometri de dues races porcines que difereixen significativament en els seus nivells de prolificitat, donant una llista de més d'un centenar de gens diferencialment expressats la funció dels quals està associada amb etapes crítiques per a la supervivència embrionària durant la gestació. Aquestes diferències d'expressió han estat validades per 12 gens que constitueixen una llista de nous candidats a exercir un paper clau en l'arquitectura genètica de caràcters relacionats amb l'eficiència reproductiva en el porc. Donat que les microRNAs (miRNAs) són coneguts reguladors post-transcripcionals de l'expressió génica, vam pensar que les diferències observades en el nivell d'expressió d'aquests gens podia respondre a un patró d'expressió de microRNAs diferent. Per validar aquesta hipòtesi, es va analitzar el perfil d'expressió de miRNAs en l'endometri de truges gestants amb nivells de prolificidad divergents, identificant 10 miRNAs madurs diferencialment expressats. Tot i que després de la seva quantificació relativa els nivells d'aquests microRNAs van resultar ser similars, es van trobar correlacions significatives entre l'expressió dels miRNAs *ssc-miR-92a* i *ssc-miR-133a* i els gens candidats *MMP8*, *PTGS2*, *PTHLH* i *SCNN1G*. A més, es va dur a terme la caracterització funcional de nou miRNAs altament implicats en reproducció identificant un total de 13 polimorfismes (SNPs) a les seves seqüències precursores. Per determinar l'efecte d'aquestes variants en l'eficiència reproductiva de les truges, es va realitzar un estudi d'associació que va revelar que el genotip per a les variants identificades a la seqüència de *ssc-mir-27a*, *ssc-mir-29b-2* i *ssc-mir-106* era determinant tant per als nivells d'expressió del miRNA madur com per als valors d'EBV. Aquests resultats suggerien que les variants genètiques a la seqüència de miRNAs precursors

juguen un paper clau en els caràcteres relacionades amb la reproducció porcina. Finalment, es va dur a terme la validació funcional de la regulació dels gens *ADM*, *HTRA3*, *PTHLH* i *VEGFA* per part dels seus microRNAs diana *ssc-miR-181d-5p*, *ssc-miR-101-3p*, *ssc-miR-144* i *ssc-miR-195-5p* respectivament, que ens va permetre establir una relació directa entre aquestes interaccions i una disminució en els seus nivells d'expressió.

INTRODUCTION

Table 1.1. Heritability (h^2) estimations for female reproductive traits. ………….……. 34

Table 1.2. Number of QTLs by pig trait class. …………………………………..…... 36

Table 1.3. Summary of significantly ($p < 0.01$) associated QTL regions and some important genes within the regions for reproductive traits in maternal pig lines. ……………..…. 37

Table 1.4. Current available NGS platforms and their characteristic features ……... 44

Table 1.5. Significant QTL for reproduction traits analyzed within the MEIBMAP project. ………………………………………………………………………………….…… 59

Table 1.6. Candidate genes analyzed within the MEIBMAP project. ………………. 62

ARTICLES AND STUDIES

Study III: Determination of polymorphisms affecting the regulatory function of reproductive-related miRNAs

Table 1. Primers used for the SNP identification by Sanger sequencing. ………………. 117

Table 2. Primers used for the mature miRNAs expression validation by RT-qPCR. …… 118

Table 3. Identified variants in the sequence of the candidate microRNAs by Sanger sequencing ……………………………………………………………………………………….. 119

Table 4. Association test results. ……………………………………………………….. 120

Study IV: Functional validation of the miRNA-mRNA interaction

Table 1. Phenotypic records of the F_2 Iberian × Meishan sows used in this study. ….… 128

Table 2. RNA-seq results for the selected genes. ……………………………………..… 129

Table 3. Primers and restriction enzyme sequences used for the 3'-UTRs cloning. ….. 130

Table 4. MiRNA targeting prediction performed by TargetScan and miRDB. ………..… 131

ANNEXES

Study I: Analysis of gene expression differences between extreme prolificacy phenotypes

Table S1. Differential expressed genes found in mRNA libraries between both extreme phenotypes. ………………………………………………………………………………… 189
Table S2. Differentially expressed genes found uniquely expressed in one of the prolificacy groups in mRNA libraries. ……………………………………………………………………… 192
Table S3. QTL mapping results for those DEGs located within at least one QTL closely related with litter size. ………………………………………………………………………………… 194
Table S4. Differential expressed miRNAs found in small RNA libraries between both extreme phenotypes. ……………………………………………………………………………….. 196
Table S5. TargetScan results showing DEG predicted as target mRNAs for our differentially expressed miRNAs. ………………………………………………………………………... 197
Table S6. Candidate novel miRNAs predicted by miRDeep in H and L small RNA libraries. ………………………………………………………………………………………………… 198

Study II: Validation of reproduction-related candidate genes

Additional file 1. Phenotypic records of the F_2 Iberian × Meishan sows used in this study. …………………………………………………………………………………………………… 199
Additional file 2. Predicted and validated expression results for the selected genes. …………………………………………………………………………………………………… 200

Additional file 3. Primers used for the genes RT-qPCR validation design. ……… 202

Study III: Determination of polymorphisms affecting the regulatory function of reproductive-related miRNAs

Supplementary Table 1. Phenotypic records of the extreme F_2 Iberian × Meishan sows used in this study. ………………………………………………………………………………. 203
Supplementary Table 2. Genotypes of the whole population for the identified SNPs. ……………………………………………………………………………………………………. 204
Supplementary Table 3. Association between the observed genotype for each variant and EBVs. ………………………………………………………………………………………………. 217
Supplementary Table 4. Mature miRNA expression analysis results. ………………….. 218

INTRODUCTION

Figure 1.1. World meat production by livestock animal from 1950-2010. …………………29

Figure 1.2. Endocrine mechanisms of the normal estrous cycle of a sow. ……................31

Figure 1.3. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. ……..………………………… 40

Figure 1.4. Basic principles from Illumina Solexa (a), Ion Torrent (b) and (c) SOLiD sequencing methodologies. ……………………………………………………………………………… 42

Figure 1.5. Overview of the RNA-seq analysis pipeline for detecting differential expression. ………………………………………………………………………………………………….. 45

Figure 1.6. RT-qPCR amplification plot. …………………………………………………….. 48

Figure 1.7. Alternative RT-qPCR methodologies to generate cDNA using stem-loop primers (A), linear miRNA-specific primers (B) or by enzymatic tailing using Poly(A) Polymerase (C) or T4 RNA Ligase (D). ………………………………………………………………………… 50

Figure 1.8. Amplicon detection by using SYBR Green (A) or TaqMan probes (B). …….. 51

Figure 1.9. Biosynthesis of miRNAs. ………………………………………………………… 53

Figure 1.10. mRNA-target recognition by microRNAs in animals. …………….………… 54

Figure 1.11. Mechanism to detect the impact of a miRNA on the regulation of a target mRNA by luciferase reporter assay. ………………………………………………………………. 56

Figure 1.12. Experimental Iberian x Meishan intercross (MEIBMAP). …………………… 58

ARTICLES AND STUDIES

Study II: Validation of reproduction-related candidate genes

Figure 1. QuantStudio™12K Flex real-time PCR gene expression results between high and low groups. ……………………………………………………………………………………… 100

Study III: Determination of polymorphisms affecting the regulatory function of reproductive-related miRNAs

Figure 1 (a-c). Association between the genotype for the identified variants and the EBVs of the pregnant sows. ………………………………………………………………………………120-21

Figure 2 (a-c). Mature miRNA expression results obtained by RT-qPCR. ………………. 122

Study IV: Functional validation of the miRNA-mRNA interaction

Figure 1 (a-d). Luciferase reporter assay results. ……………………………….…133

Figure 2 (a-d). Summary of the previous endometrial transcriptome analysis results. … 134

Córdoba S., Balcells I., Castelló A., Ovilo C., Noguera J.L., Timoneda. O, Sánchez A. *Endometrial gene expression profile from pregnant sows with extreme phenotypes for reproductive efficiency*. Scientific Reports (5): doi:10.1038/srep14416

Córdoba-Terreros S., Balcells I., Castelló A., Ovilo C., Noguera J.L., Timoneda. O, Sánchez A. *Genetic polymorphisms in miRNAs are associated with porcine extreme phenotypes for reproductive efficiency*. Reproduction (under review)

Córdoba-Terreros S., Balcells I., Castelló A., Ovilo C., Noguera J.L., Timoneda. O, Sánchez A. *Key genes for litter size control show significant expression differences in the endometrium of pregnant sows with extreme phenotypes for reproductive efficiency*. Animal Genetics (Submitted)

Córdoba S., Dhorne-Pollet S., Castelló A., Balcells I., Ovilo C., Noguera J.L., Timoneda O., Sánchez A. *Expression of candidate genes affecting porcine reproductive efficiency is regulated by microRNAs ssc-miR-101-3p, ssc-miR-144-3p and ssc-miR-195-5p.* (Manuscript in preparation)

Núñez-Hernández, Fernando; Pérez, Lester J; Muñoz, Marta; Vera, Gonzalo; Tomás, Anna et al. (2015) *Identification of microRNAs in PCV2 subclinically infected pigs by high throughput sequencing*. Veterinary research - vol. 46 (1) p. 18

Núñez-Hernández, Fernando; Pérez, Lester J; Vera, Gonzalo; Córdoba, Sarai; Segalés, Joaquim et al. (2015) *Evaluation of the capability of the PCV2 genome to encode miRNAs: lack of viral miRNA expression in an experimental infection*. Veterinary research - vol. 46 (1) p. 48

Oriol Timoneda, Fernando Núñez, Ingrid Balcells, Marta Muñoz, Anna Castelló, Gonzalo Vera, Lester J. Pérez, Raquel Egea, Gisela Mir, Sarai Córdoba, Rosa Rosell, Joaquim Segalés, Anna Tomàs, Armand Sánchez, José I. Núñez (2013) *The role of viral and host microRNAs in the Aujeszky's disease virus during the infection process*. PLoS ONE 9(1): e86965. doi:10.1371/journal.pone.0086965

Timoneda O, Balcells I, Córdoba S, Castelló A, Sánchez A (2012) *Determination of Reference microRNAs for Relative Quantification in Porcine Tissues*. PLoS ONE 7(9): e44413.doi:10.1371/journal.pone.0044413

ADM	Adrenomedullin
AFLP	Amplified fragment length polymorphism
BLUP	Best linear unbiased predictors
Bp	Base pair
cDNA	Complementary DNA
CES1	Carboxylesterase 1
CL	Corpora lutea
Ct	Threshold cycle
CV	Coefficient of variation
DD	Differential display
DE	Differential expression
DEG	Differentially expressed gene
DGCR8	DiGeorge Syndrome Critical Region 8
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
EBV	Estimated Breeding Value
ECM	Extracellular matrix
endo-siRNAs	Endogenous short interfering RNAs
EPOR	Erythropoietin receptor
eQTL	Expression quantitative trait loci
ESR	Estrogen receptor
ESTs	Expressed sequence tags
FAO	Food and Agriculture Organization of the United Nations
FC	Fold change
FDR	False discovery rate
FOXA	Forkhead transcription factor 2.
FSH	Follicle stimulating hormone
FSHβ	Follicular-stimulating hormone beta subunit
Fw	Forward
FXYD3	FXYD Domain Containing Ion Transport Regulator 3

GF	Growth factor
GL	Gestation length
GL	Gestation length
GnRH	Gonadotropin releasing hormone
GO	Gene ontology
H	High
Hsa	Homo sapiens
HT-NGS	High Troughput Next Generation Sequencing
HTS	High Troughput Sequencing
Ib	Iberian porcine breed
ILK-1β	Interleukin 1 β
IPA	Ingenuity Pathway Analysis
KASP	Kompetitive Allele Specific PCR
KLF5	Kruppel-like factor 5
KLK1	Tissue Kallikrein gene
L	Low
LH	Luteinizing Hormone
lncRNAs	Long non-coding RNAs
Log2	Logarithm base 2
MAGRAMA	Ministerio de Agricultura, Alimentación y Medio Ambiente
Me	Meishan porcine breed
miRISC	MicoRNA-mediated Silencing Complex
miRNA	Small non-coding RNA molecule, micro RNA
miRNA*	miRNA star
MMP8	Matrix metalloproteinase
mRNA	Messenger RNA
MUM	Mummified foetuses at birth
ncRNAs	Non-coding RNAs
NF	Number of fetuses
NGS	Next Generation Sequencing

NSB	Number of Stillborn Piglets
Nt	Nucleotide
OR	Ovulation rate
OR	Odds Ratio
PAP	Poly A polymerase
PCR	Polymerase Chain Reaction
PGs	Prostaglandins
PION	Gamma-secretase activation protein
piRNAs	Piwi-protein-interacting RNAs
pre-miRNA	Precursor microRNA
pri-miRNA	Primary microRNA
PRL	Prolactin
PRLR	Prolactin receptor
PTGS2	prostaglandin G/H synthase or cyclooxygenase COX2)
PTHLH	parathyroid hormone-like hormone
qPCR	Quantitative Polymerase Chain Reaction
QTL	Quantitative Trait Loci
RBPs	RNA Binding Proteins
RIN	RNA Integrity Number
RNA	Ribonucleic Acid
Rnase III	Ribonuclease III Enzyme
RNA-seq	RNA Sequencing
RPKM	Reads Per Kilobase of Exon Model Per Million Mapped Reads
RQ	Relative Quantity
Rrna	Ribosomal RNA
RT	Reverse Transcription
-RT	Minus Reverse Transcription
RT-qPCR	Reverse Transcriptase Real Time Quantitative Polymerase Chain Reaction
Rv	Reverse
SAGE	Serial Analysis of Gene Expression

SCNN1G	Sodium Channel, Non Voltage Gated 1 Gamma Subunit
SD	Standard deviation
SDCBP-2	Syndecan Binding Protein (Syntenin) 2
SE	Standard error
SGSC	Swine Genome Sequencing Consortium
SNP	Single Nucleotide Polymorphism
Ssc	Sus scrofa
SSCn	Sus scrofa Chromosome (where n corresponds to the chromosome number)
TBA	Total number of piglets born alive
TF	Transcription factor
Tm	Melting Temperature
TN	Teat number
TNB	Total number of piglets born
TNF	Tumor necrosis factor
USDA	United States Department of Agriculture
VEP	Variant effect predictor

1. GENERAL INTRODUCTION

1.1. PIG PRODUCTION

The pig (*Sus scrofa*) constitutes economically one of the most important species in livestock. Porcine is the most widely eaten meat in the world accounting for over 36% of the world meat intake (FAO 2014) (Figure 1.1). Its production has increased every year representing an increment of a 4.38% since 2014 (MAGRAMA 2014), placing Spain at the 4th position among the 20 highest producing countries of porcine meat. This upward trend evidences the economic importance of this sector, especially in Catalonia which represents a 43.1% of the total meat produced in Spain with 1,551,166 tons produced in 2014 (IDESCAT 2015).

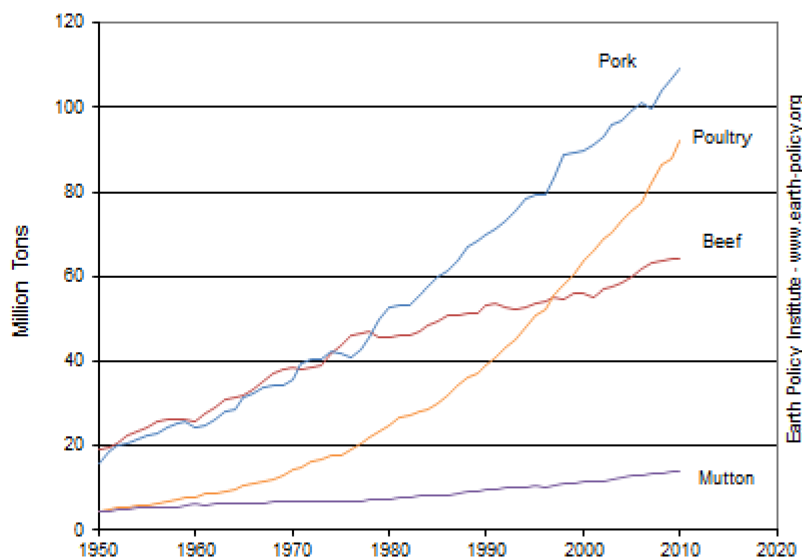


Figure 1.1. World meat production by livestock animal from 1950-2010 (FAO, 2014)

Although pigs were one of the first species to be domesticated approximately 9000 years ago, development of pig production did not started until 1960 when the first group of exotic pig breeds (Large Whites, Tamworth and Berkshire) were imported by the Department of Livestock Development from the United Kingdom (Larson et al. 2010). Before the introduction of these exotic breeds farmers had to rely on the relatively slow growing of native pigs; however, imported pigs were soon used for breeding improvement and it was throughout the 1960s and 1970s when producers started to raise crossbred pigs as a source of income (Groenen et al. 2012).

1.1.1. Improvement of pig production

The reproductive performance of a sow is one of the key factors affecting production profitability in pigs (Onteru et al. 2009). Thus, from an economic point of view, reproductive efficiency is one of the most important factors in livestock. To date, quantitative approaches used in animal breeding relied upon recording data from phenotypic traits of interest on a large number of individuals (Goddard & Hayes 2009). These records were analyzed with a wide set of statistical methods that allowed the identification and selection of superior individuals as the parents of the next generation (Hill 2014). This strategy became highly efficient when dealing with traits that present moderately or highly heritability; however, because of its complex genetic architecture, selection and improvement of reproductive-related traits has been rather challenging (Bidanel 1993).

1.2. REPRODUCTION IN SWINE

Unlike other livestock species, the pig is a multiparous specie that gives birth to a large number of offspring at the same time (Bidanel 2015). Sows reach their sexual maturity at 5-7 months old, as a result of the interaction of internal (genotype, breed…) and external (nutrition, health, environment…) factors. From birth sows present all their primary follicles in both ovaries (approx. 400,000), but it is after puberty when the first fertile estrous cycle occurs (Hughes & Varley 1980). Their estrous cycle classifies this specie as a continuous polyestrous, which means that reproduction has no seasonality and regular cycles are repeated throughout the year, every 21 days except during pregnancy and lactation (WELLS 1946).

1.2.1. Reproductive cycle of a sow

The sexual cycle of a sow spans a period of 18–24 days and it is regulated by changes in the levels of circulating hormones determined by the hypothalamic-pituitary-ovarian axis (Rydhmer 2000). The hypothalamus, located at the base of the brain, secretes the gonadotropin releasing hormone (GnRH) which regulates the anterior pituitary gland, modulating the blood levels of the follicle stimulating hormone (FSH) and the luteinizing hormone (LH) (Krzymowski & Stefańczyk-Krzymowska 2008). These two hormones stimulate the production of the two ovarian hormones estrogen and progesterone, which in turn regulate the whole reproductive process. Sow's

reproductive cycle consists of a follicular phase of 5–7 days and a luteal phase of 13–15 days (Figure 1.2).

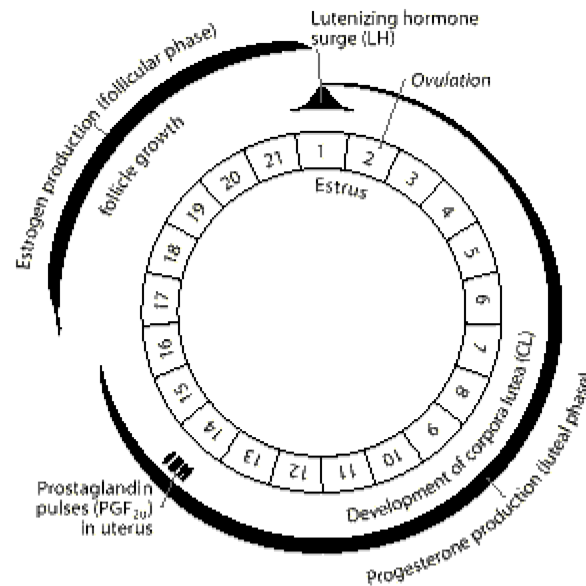


Figure 1.2. Endocrine mechanisms of the normal estrous cycle of a sow (modified from Roy N. Kirkwood *et al*. 1997).

1.2.1.1. Follicular phase

It comprises from the end of the luteal phase (day 16 of oestrus cycle) to the ovulation of the follicles. During this phase, small ovarian follicles develop into large, pre-ovulatory follicles (Rydhmer 2000). The number of oocytes released by both ovaries in an estrous (ovulation rate) is between 15 to 30 follicles, depending on age, nutritional status and other factors. Ovulation is spontaneous and occurs during the second half of the estrus, around 38-42 hours after its onset, and lasts around 1-6 hours (Johnson et al. 1999). This period is characterized by the secretion of estradiol and LH hormone, which are essential for the development of these follicles (Krzymowski & Stefańczyk-Krzymowska 2008).

1.2.1.2. Luteal phase

Once an oocyte is released from a ruptured follicle, the remaining cells within the ruptured follicle continue to develop and form the *corpus luteum* (CL). The luteal phase appears with the development of several *corpus luteum*, collectively called *corpora lutea*. The *corpora lutea* secrete progesterone, which blocks the secretion of both, FSH and LH hormones (Bertoldo et

al. 2012). If there is no fertilization, and oocytes are not fertilized, the uterus starts to secrete prostaglandin that induces the regression of the *corpora lutea* (luteolysis) which end the secretion of progesterone. In absence of progesterone a new estrus cycle begins (Sun & Nagai 2003). If there is fertilization, the oocytes become viable embryos and attach to the uterus. At approximately day 11-12, the attached embryo secretes estradiol that redirects uterine prostaglandins from the vascular system to the lumen of the uterus blocking the regression of the *corpora lutea*. Then, there is a feedback to the *corpora lutea* that signals them to continue to produce progesterone preparing the uterus for pregnancy establishment (Rydhmer 2000).

1.2.1.3. Pregnancy establishment

Pigs tend to produce large litters in a relatively short period of time. The establishment of pregnancy begins about 11-12 days after the beginning of oestrus and is a process that lasts an average of 114 days (3 months, 3 weeks and 3 days), although it can vary with each pregnancy (Bazer 2013). In pigs, this process comprises three main periods: post-conception period (days 1–10 of pregnancy), maternal recognition of pregnancy (days 11–13) and implantation (days 14–19):

- **Pre and Post conception period**

The post-conception period starts with fertilization and lasts until day 10 after oestrus. During this period, CL is developed. The main product of CL is progesterone, which induces the preparation of endometrium for implantation (Mathew et al. 2011). Moreover, it is thought that during this period uterine immune system is activated. After entering the female reproductive tract, gametes and other components of boar's seminal plasma are recognized by the sows oviduct and the uterine horns (Kaczmarek et al. 2010). Components of boar semen induce the infiltration of leucocytes into the uterine lumen, clearing the uterus of redundant spermatozoa and microorganisms introduced at mating, thus promoting and regulating the local immune responses (O'Leary et al. 2004). It has been suggested that this interaction of the boar semen with the reproductive tract of the sow, may increase embryo survival and litter size.

- **Maternal recognition of pregnancy**

Maternal recognition of pregnancy is the process in which embryo signals its presence in the uterus and the sow receives and accepts this signal (Bazer 2013). The pig conceptus (embryo and its associated extra-embryonic membranes) secretes abundant amounts of estrogens that act on the uterus and/or corpus luteum to ensure their maintenance for the progesterone production, the required hormone for pregnancy in most mammals (Mathew et al. 2011).

- **Implantation period**

Succeeding the embryonic signal for the maternal recognition of pregnancy, porcine embryos remain free-floating until days 13–14 of pregnancy, when they attach to the uterus. As the fetus reaches maturation and the conclusion of pregnancy is needed, the uterus secretes prostaglandin causing parturition (Dey et al. 2004).

Recent transcriptomic analysis have revealed that several genes involved in developmental processes such as transporter activity, calcium ion binding, lipid metabolic processes, hormone activity, cell motility and apoptosis are differentially expressed between pregnant and cyclic pigs on day 14 after ovulation (Chen et al. 2014; Samborski et al. 2013; Østrup et al. 2010). Thus, indicating that dramatic changes take place during conceptus implantation.

1.3. GENETIC BASIS OF SWINE REPRODUCTION

Although there is a wide diversity of reproductive strategies, basic principles involved in sexual reproduction are relatively conserved and well defined among mammals (Furnes & Schimenti 2007). However, the underlying molecular and genetic basis of each step involving these general processes still remains unknown. At first, genetic improvement of reproductive traits was mainly focused on phenotypic selection rather than using genotypic information (Spötter & Distl 2006a). As previously discussed, the use of these traditional selection methods has not been successful in most livestock species, due to the complex genetic basis and the low heritability (Bidanel 2015) that these traits present (Table 1.1).

Table 1.1. Heritability (h^2) estimations for female reproductive traits (Bidanel *et al.*, 2011).

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

These limitations have led to a growing interest in the identification of specific genes and genomic regions involved in the variability and regulation of reproductive traits. At present, researchers have focused in the identification of genes or genomic regions influencing reproductive phenotypes, trying to understand the genetic control of female reproduction in order to develop a more efficient selection of the candidates for reproductive efficiencies (Wilkie et al. 1999; Rohrer et al. 1999; Du et al. 2014; King et al. 2003).

1.3.1. Litter size

Since the main objective of the pig industry has been to obtain the highest number of piglets weaned per unit time at the lowest cost, reproductive traits related with litter size and pre-weaning viability have become one of the most relevant traits from a genetic and economic point of view (Rothschild 1996). Improvements in litter size across the swine industry have occurred through different selection schemes such as phenotypic, family index, hyper-prolific-based selection or best linear unbiased prediction method (BLUP) (Spötter & Distl 2006a), which allows to estimate a breeding value (EBV) for each animal that directly correlates with its suitableness to be selected as a reproducer (Hill 2014).

One of the main determinants of litter size is failure of the developing foetus to survive (Spötter & Distl 2006b). Despite it is difficult to determine the exact events and functions involved in pregnancy success or failure, the distinct components affecting this trait have yet been determined: ovulation rate, embryonic development, uterus capacity and particularly foetal survival and pre-weaning losses which are the most important component traits used in swine

breeding programmes (Rydhmer 2000) for their direct impact in the final litter of a sow. These losses can occur at each stage of development, but the main critical stages are early (days 10–30 of gestation) and mid-gestation (days 50–70 of gestation), which represent around 20–30% and 10–15% of embryonic loss respectively (Spötter & Distl 2006b) and are primarily determined by the uterine capacity of the pregnant sows (Vallet & Freking 2007; Ford et al. 2001). In pigs, litter size is estimated through the total number of piglets born (TNB) and the number of piglets that born alive (NBA) per parity. Considering that the TNB is the sum of the NBA and the number of stillborn piglets (NSB), the final litter of a sow can be determined by this formula:

$$\mathbf{TNB = NBA + NSB}$$

When performing the genetic study of reproductive traits, the success of the main stages involved in reproduction is estimated as diverse phenotypic records, which include endocrine measures (hormone levels), morphologic measures of reproductive organs (teat number, length and placement, uterine capacity and length), fertility related traits (fertilization rate, ovulation rate) and litter measures (embryo survival, counts of live, dead, mummified, weaned descendants) and other general reproductive traits as age at puberty or gestation length (Vanderhaeghe et al. 2013). The combination of all these phenotypic records with genotypic information would greatly improve the final litter of a sow (Distl 2007).

1.3.2. Reproductive QTLs

Quantitative trait loci (QTL) are defined as genomic regions which contain one or more genes that affect the variation of a quantitative trait (Andersson 2001). Over the past years, advances in the porcine genetic linkage map have allowed the identification of thousands of quantitative trait loci (QTL) for a wide range of economically important phenotypes in pigs such as growth and body composition, carcass and meat quality, reproduction, and disease resistance (Cassady et al. 2001; de Koning et al. 2001; Rathje et al. 1997). Initially, most of the QTL experiments performed to determine those regions underlying relevant traits to the pig industry were carried out by using initial linkage maps based on recombination frequency. These early QTL scans used around 300 to 700 pigs, usually coming from an F_2 obtained by generally crossing European Wild Boar with a commercial breed or crossing the exotic Chinese Meishan

breed with a commercial breed (Rothschild et al. 2007; Campbell et al. 2003; Buske et al. 2006a). The first QTL discovered based on this methodology, was a major locus for fat deposition on porcine chromosome 4 (Andersson et al. 1994). Later on, researchers tend to originate these F_2 families using phenotipically divergent commercial breeds or large commercial synthetic lines. But to enhance and improve current selection procedures, it is necessary to identify reliable markers.

Rapidly evolving genomics technologies and the recent use of high-density single nucleotide polymorphism (SNP) panels to perform large scale SNP association analyses have extended beyond experimental intercrosses to outbred populations resulting in higher resolution QTL mapping and increasing the number of discovered QTLs and eQTL(Ernst & Steibel 2013). The number of mapped QTLs is still growing and most of the updated QTL mapping results are available through the well-developed QTL database called PigQTLdb (Hu et al. 2013). A summary of the current knowledge regarding porcine QTLs can be shown at Table 1.2.

Table 1.2. Number of identified QTLs by pig trait class. (Extrected from PigQTLdb, release 27 - August 2015; http://www.animalgenome.org/QTLdb/)

Trait Class	Number of QTL
Meat & Carcass Quality	7,277
Health	2,061
Production	1,424
Reproduction	1,235
Exterior	1,034

Because of the relevance of reproductive traits, large scale QTL and candidate gene studies have been conducted to discover potential markers that are actively incorporated by the pig industry in marker-assisted selection schemes. To date, the most significant QTLs associated with porcine reproductive traits that have been identified are: SSC3, SSC8, SSC9, SSC10 and SSC15 for ovulation rates (Rathje et al. 1997; Rohrer et al. 1996; Wilkie et al. 1999; Campbell et al. 2003), SSC7, SSC8, SSC12, SSC13, SSC14 and SSC17 for total number piglets born (de Koning et al. 2001; King et al. 2003; Noguera et al. 2009), SSC4 and SSC13 for number of stillborn (Wilkie et al. 1999; Cassady et al. 2001) and SSC8 for uterine capacity and prenatal survival (Rohrer et al. 1999; King et al. 2003). QTL discovery in pigs has advanced rapidly and currently several eQTL experiments are underway. Although in some cases there is a lack of

similar markers and slightly different trait definitions and measurements, these new methodologies will help to improve efficiency of pig production and in general, make pigs a more useful biomedical model (Rothschild et al. 2007).

1.3.3. Candidate genes for litter size in pigs

The final goal of complex traits dissection is to identify the involved genes and to decipher their cellular roles and functions. Although more than six hundred QTLs for litter size have been identified, a limited number of useful genes have been found to have significant associations with reproductive traits (Buske et al. 2006b; Zhou et al. 2009). Despite some major candidate genes such as estrogen receptor (*ESR*), prolactin receptor (*PRLR*), follicular-stimulating hormone beta subunit (*FSHβ*), erythropoietin receptor (*EPOR*), osteopontin (*OPN*) and prolactin (*PRL*) have been identified to play a key role in sows' reproductive efficiency (Hu et al. 2013) (Table 1.3), true causal genes responsible for this trait still remain scant due to the large disequilibrium linkage blocks present in the genome of livestock species(Rothschild 1996).

Table 1.3. Summary of significantly associated QTL regions and some important genes within the regions for reproductive traits in maternal pig lines (modified from Onteru SK, et al. 2011).

Trait	Nºof QTL regions	SSC for candidate regions	Relevant genes in the QTL regions (SSC)
TNB1	14	2, 3, 4, 7, 8, 14, 16	*MEF2C*, *RASA1*, *HTR1A*
TNB2	33	3, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17	*PLSCR4*, *PLSCR5*, *PTX3*, *SEC23B*
TNB3	28	1, 2, 3, 4, 6, 8, 9, 12, 13, 14, 15, 18, X	*BCL7B*, *NIPAL2*, A Novel protein (15)
NBA1	11	1, 2, 3, 4, 12, 14, 16	*IGFBPL1*, *RASA1*, *MEF2C*, *HTR1A*
NBA2	22	1, 5, 7, 10, 11, 12, 13, 14	*PLSCR4*, *PLSCR5*, ATGR1, *TBX3*
NBA3	9	2, 3, 4, 6, 12, 15	*BCL7B*, *ROR1*, A Novel protein (15)
SB1	25	2, 4, 6, 8, 9, 10, 12, 15, 16, 17, 18	*EYA3*, *RPLPO*, *HNRNPD*
SB2	17	1, 3, 4, 5, 6, 8, 10, 14, 16	*CDH20*, *SS18*, *TAF4B*, *KCTD1*
SB3	21	1, 2, 3, 4, 5, 6, 8, 12, 13, 15, 17, 18	*FGGY*, *RELL1*, *ACCN1*
MUM1	37	1, 2, 4, 6, 9, 10, 13, 14, 15, 17	*ESR1*, *AHR*, *AQP7*
MUM2	26	1, 2, 3, 4, 5, 8, 9, 10, 13, 14, 15, 16, 17	*EEA1*, *ACAD11*, *NPHP3*, *CCRL1*, *USB5*
MUM3	41	1, 2, 3, 4, 6, 7, 9, 10, 11, 13, 14, 15, 16, 17, 18, X	*ECDHE2*, *HSPH1*, *CD96*, *ZEBD2*
GL1	21	2, 4, 5, 6, 9, 11, 13, 15, 16	*FSHB*, *CRSP2*, *CALCA*, *PTH*
GL2	12	3, 6, 7, 9, 10, 11, 17	*MATN3*, *EPS15*, *MAFB*
GL3	20	1, 6, 7, 9, 13, 14 18	*FGF7*, *CHGA*, *VEGFA*

TNB, total number born; NBA, number born alive; SB, number of stillborn; MUM, mummified foetuses at birth; GL, gestation length; 1, 2 and 3 represent parity 1, 2 and 3, respectively

In order to perform a wider application across the different pig breeds and lines, these two functional genomic approaches have to be merged.

1.4. PIG GENOMICS

Molecular genetics and genomics are the fields of biology that study the structure and function of genes and genomes at a molecular (nucleic acids) level. Since the first eukaryote genome was sequenced in 1997 (Mewes et al. 1997), and the first draft sequences of the human genome were published in 2001 (Cheung et al. 2001), the need to decipher the genetic basis of economically important production traits in pigs led to the development of several methodologies for retrieving structural and functional genome information: development of genetic markers (Davies et al. 1994; Coppieters et al. 1995; Groenen et al. 1995), establishment of genetic linkage (Archibald et al. 1995) and cytogenetic maps (Echard et al. 1992; Yerle et al. 1995) and identification of QTLs. In the early 1990s, the first coordinated efforts to understand the pig genome were initiated with the development of the international PiGMaP gene mapping project (Yerle et al. 1995; Archibald et al. 1995). Later on, the projects initiated by the USDA and the US agricultural universities made possible the publication of two significant linkage maps, the largest containing over 1,200 microsatellite markers. Since that, new gene markers such as microsatellites, amplified fragment length polymorphism (AFLPs), and single nucleotide polymorphisms (SNPs) have been continuously identified and mapped (Ernst & Steibel 2013). Although these approaches allowed a rapid and low-cost study of the genotypes of a large number of individuals, what greatly contributed to characterize pigs at the molecular level was the formation in 2003 of The Swine Genome Sequencing Consortium (SGSC) (Schook et al. 2005). This consortium was created by academic, government and industry representatives for sequencing the pig genome and nine years after its foundation, in November 2012 the first assembly of a domestic pig genome sequence was published (Groenen et al. 2012).

The availability of a pig genome and the ability to generate genome-scale data sets associated to high throughput sequencing techniques such as transcriptome analyses of different reproductive tissues have allowed the identification and characterization of markers, pathways

and genes responsible for the genetic complexity of reproductive traits (Onteru et al. 2009). To date, main used strategies to detect those genes affecting litter size and its components have been: linkage analyses based on the identification of genomic regions linked with a phenotypic reproduction trait and more recently, genome-wide gene expression profiling, that has become a successful strategy for identifying a higher number of candidate genes related to reproduction in livestock (Du et al. 2014; Esteve-Codina et al. 2011; Ross et al. 2009; Sun et al. 2011).

1.4.1. Transcriptome profiling

In recent years, transcriptome characterization has seen a remarkable rise, becoming a hot topic in genomic research either in human or animal genetics (Tuggle et al. 2007). The knowledge obtained by deciphering the pig genome and advances in molecular genetics, such as the transcriptomic analysis by RNA sequencing, have provided a powerful tool to better understand the genetic architecture of prolificacy-related traits. The use of microarrays and large-scale transcriptome analysis to identify differentially expressed genes in specific tissues, cell types or breeds has shed light on many aspects of porcine production traits (Samborski et al. 2013; Bauersachs & Wolf 2012; Franczak et al. 2013; Liu et al. 2014; Siqueira et al. 2014; Ramayo-Caldas et al. 2012; Corominas et al. 2013; Puig-Oliveras et al. 2014). Despite this, there have only been a few comparative studies on uterine function for prolific pigs and a low number of experiments regarding differences in endometrial gene expression between porcine breeds have yet been performed (Gu et al. 2014; H. Zhang et al. 2013; Zhou et al. 2009).

1.4.1.1. High-throughput sequencing (RNA sequencing)

Current research in biology, biotechnology, and medicine requires fast genome and transcriptome analysis technologies (Mutz et al. 2013). Whole-transcriptome shotgun sequencing also known as RNA-seq, is a recently developed approach that uses high-throughput sequencing technology for characterizing the RNA content and composition of a given sample (Morin et al. 2008). Until the arrival of RNA-seq, microarrays were the standard tool for gene expression quantification. Although both techniques are generally in good agreement regarding relative gene expression quantification (Nookaew et al. 2012), microarray technology is limited towards the amount of RNA, the quantification of transcript levels and the

sequence information. The main technological limitation of RNA-seq experiments is that sequence information from transcripts cannot be retrieved as a whole (Wang et al. 2009). To solve this, once the RNA has been converted into cDNA and sequenced on a high-throughput platform generating millions of short (25 to 300 bp) reads, transcripts are randomly decomposed into short reads of several hundred base pairs. If there is no reference genome or transcriptome information, it is necessary to first reconstruct transcripts from these short reads (or read pairs), which is called "de novo" assembly (Grabherr et al. 2011). If on the contrary, we have a transcript or genome information available; reads can be directly aligned onto this reference (Figure 1.3).

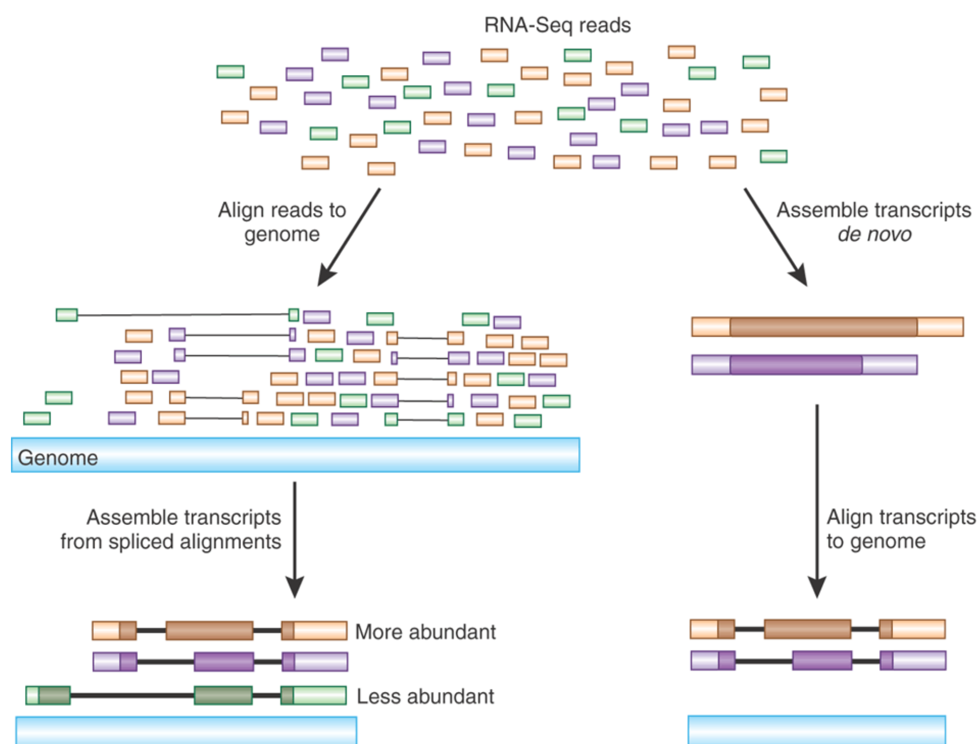


Figure 1.3. Transcript assembly and quantification by RNA-Seq (Modified from Haas BJ *et al.*, 2010).

The currently available high-throughput next generation sequencing (HT-NGS) platforms differ substantially in their chemistry and processing steps (Table 1.4) and can be classified into three main groups: first generation, second generation and third generation HT-NGS platforms (Mihály & Gyorffy 2011).

- **First generation sequencers**

These first automatic sequencers used fluorescently labeled dideoxynucleotides that were analyzed on a capillary electrophoresis to produce a chromatogram or electropherogram, from which the sequence was then deduced on the computer (Karger & Guttman 2009). This technology allowed to sequence up to 96 DNA samples at the same time in only a few hours. The length of the generated sequences was around 500-1000 bases. The increment on the length of the reads compared to the manual processing of the Sanger-sequencing, along with the development of new strategies for large-scale sequencing (Whole Genome Shotgun Sequencing) increasingly facilitated assembly of genomic sequences (Buermans & den Dunnen 2014). Examples of these first generation sequencers included the ABI Prism from Applied Biosystems and the CEQ-serie from Beckman Coulter.

- **Second generation sequencers**

After the first draft of the human genome (3,000 million nt.) was published, in 2001 which cost nearly 3,000 million dollars (1$/nt), encouraged scientists to look for cheaper solutions as sequencing costs were unaffordable for any laboratory (van Dijk et al. 2014). In this context, the called second generation sequencers were developed, which were able to generate hundreds of thousands of sequences reactions in parallel (high-throughput) by immobilizing these reactions into a solid surface. Thus, the quantity of reagents required is minimized and the cost per read base decreases (Pareek et al. 2011). Some of the most widely known sequencers belonging to this group are the GS-FLX (454) from Roche, the ABI SOLiD form Applied Biosystems, the Genome Analyzer from Illumina and Ion (PGM, Proton, S5) from Ion Torrent.

The GS-FLX (454) was based on the DNA pyrosequencing. It is a non-fluorescent technique that measures the release of pyrophosphate in a polymerization reaction using a series of coupled enzymatic reactions that emit light whenever a nucleotide is incorporated (Margulies et al. 2005). This emission produces an image that is subsequently analyzed and interpreted by the computer, sending back the whole nucleotide sequences. At the same time, two other companies developed other technologies for massively parallel sequencing DNA. Solexa-Ilumina, which sequencing technology was based in a DNA polymerization where a fluorescently labeled nucleotide was incorporated (Bentley et al. 2008). In this method, to

determine the sequence, four types of reversible terminator bases (RT-bases) are added and non-incorporated nucleotides are washed away (Figure 1.4-a). Then, a camera takes images of the fluorescently labeled nucleotides and the dye along with the terminal 3' blocker, is chemically removed from the DNA, allowing for the next cycle to begin. Unlike pyrosequencing, the DNA chains are extended one nucleotide at a time and image acquisition can be performed at a delayed moment. And the SOLiD (Sequencing by Oligonucleotide Ligation and Detection) method from Applied Biosystems, which is based in a sequence by ligation to the DNA chain of labeled octamers with a known sequence (McKernan et al. 2009). Subsequently, the fluorescent signal emitted after each ligation is detected (Figure 1.4-c).

Ion Torrent Systems Inc. developed a totally different technology based on the detection of those hydrogen ions that are released during the DNA polymerization (Rothberg et al. 2011). A micro well containing the template DNA strand is flooded with each type of single nucleotides. If the introduced nucleotide is complementary to the leading template, it is incorporated into the growing complementary strand causing the release of a hydrogen ion that activates a hypersensitive ion sensor. As the number of released hydrogens is proportional to the electronic signal, the sequence pattern can be predicted (Figure 1.4-b).

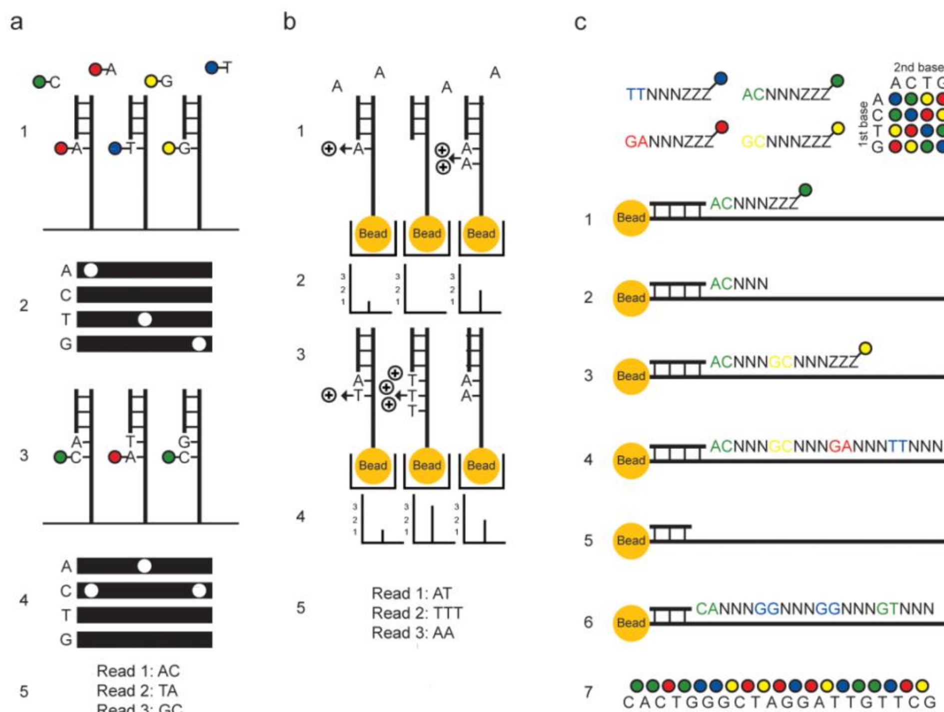


Figure 1.4. Basic principles from Illumina/Solexa (a), Ion Torrent (b) and (c) SOLiD sequencing technologies (Corney *et al.*, 2015)

Thanks to the development of these second-generation sequencers, the final cost of each nucleotide decreased from 10$ in 1990 to 0.01$ in 2005 (Pareek et al. 2011).

- **Third generation sequencers**

The continuous need to descend sequencing costs and increase the reliability of the resulting sequences has led to the recent development of the called third-generation sequencers (Pareek et al. 2011). This new era is based on the sequencing of a single DNA molecule (single molecule real-time sequencing). The first third generation sequencer designed by Helicos BioSciences, was able to real time sequence billions of small unique DNA molecules attached to a solid surface, generating fragments of around 25-45 bases (Harris et al. 2008). In a step further, the companies Pacific Biosciences and Oxford Nanopore, developed a technology able to read up to 1000 nucleotides in a single run, solving all the problems associated with second-generation sequencers (ie, homopolymer regions, tandem repeats…). This new technology, is a completely different approach where the DNA polymerase its anchored to a solid surface called “nanopore” (Pennisi 2012). Finally, ZS Genetics is using electron microscopy to read the DNA sequence tagged with iodine, bromine or trichloromethyl, directly on an electronic image. Some examples of this type of sequencers are Helicos tSMS, MinION, gridION, and ZX Genetics.

- **Fourth generation sequencers**

Althought this methodology is still very experimental, fourth generation sequencers would be able to carry out a sequencing experiment on individual cells in a histological section, i.e. in their own biological context. Applications of this new methodology would be the interrogation of those DNA sequences likely to have undergone somatic mutations, differentiate between members of a gene family and perform multiplex detection of transcripts (Koboldt et al. 2013).

Table 1.4. Current available NGS platforms and their characteristic features (adapted from Buermans *et al.*,2014)

	Sequencer	Sequence by	Detection	Run types	Read length (bp)	Reads per run	Output per run
Roche	GS FLX Titanium XL +	Synthesis	Pyrophosphate detection	Single end	700	1 million	700 Mb
	GS Junior System	Synthesis	Pyrophosphate detection	Single end	400	0.1 million	40 Mb
Life Technologies	Ion torrent	Synthesis	Proton release	Single end	200–400	4 million	1.5–2 Gb
	Proton	Synthesis	Proton release	Single end	125	60–80 million	8–10 Gb
	Abi/solid	Ligation	Fluorescence detection of di-base probes	Single & paired-end	75 + 35	2.7 billion	300 Gb
Illumina /solexa	HiSeq2000/ 2500	Synthesis	Fluorescence; reversible terminators	Single & paired-end	2 × 100	3 billion	600 Gb
	MiSeq	Synthesis	Fluorescence; reversible terminators	Single & paired-end	2 × 300	25 million	15 Gb
Pacific biosciences	RSII	Single molecule synthesis	Fluorescence; terminally phospholinked	Single end	50% of reads > 10 kb	0.8 million	5 Gb
Helicos	Heliscope	Single molecule synthesis	Fluorescence; virtual terminator	Single end	~ 30	500 million	15 Gb

Once the RNA sequencing has been performed, making sense of the huge amount of data generated depends on the scientific question of interest (Oshlack et al. 2010). If the aim of our study is, for example, determining differences in allele-specific expression, we would require a precise estimation of the prevalence of transcribed single nucleotide polymorphisms (SNPs). If we are identifying fusion genes or aberrations in cancer samples, it would be necessary to find novel transcripts and/or RNA editing events (Rapaport et al. 2013). However, the primary use of RNA-seq is gene expression profiling between samples. In this case, it is necessary to count the reads that fall onto a given transcript, which serves as a digital measurement of transcript abundance being the starting point for gene expression quantification (Robinson & Oshlack 2010). The constant development of new software and the numerous characteristics of each transcriptomic analysis (selected specie, sequencing technology, quality of the reference genome...) have increased the number of available pipelines to analyze these massive sequencing data, being almost exclusively for each study (Carvalho & Rustici 2013). The standard and most common pipeline for detecting differential expression (DE) in RNA-seq consists in five main steps (Figure 1.5): mapping, assembly, data normalization, statistical test of DE and biological contextualization of the obtained results (R. Huang et al. 2011). First,

reads are mapped to the genome or transcriptome. Subsequently, mapped reads are assembled into gene, exon or transcript-level expression, depending on the aims of the study. After so, summarized data has to be normalized in order to perform statistical testing for differential expression (DE). Then, statistical analysis leads to a ranked list of genes with associated *p-values* and fold changes. Finally, to gain biological insight from these, systems biology approaches should be performed, similar to those performed on microarray experiments (Oshlack et al. 2010).

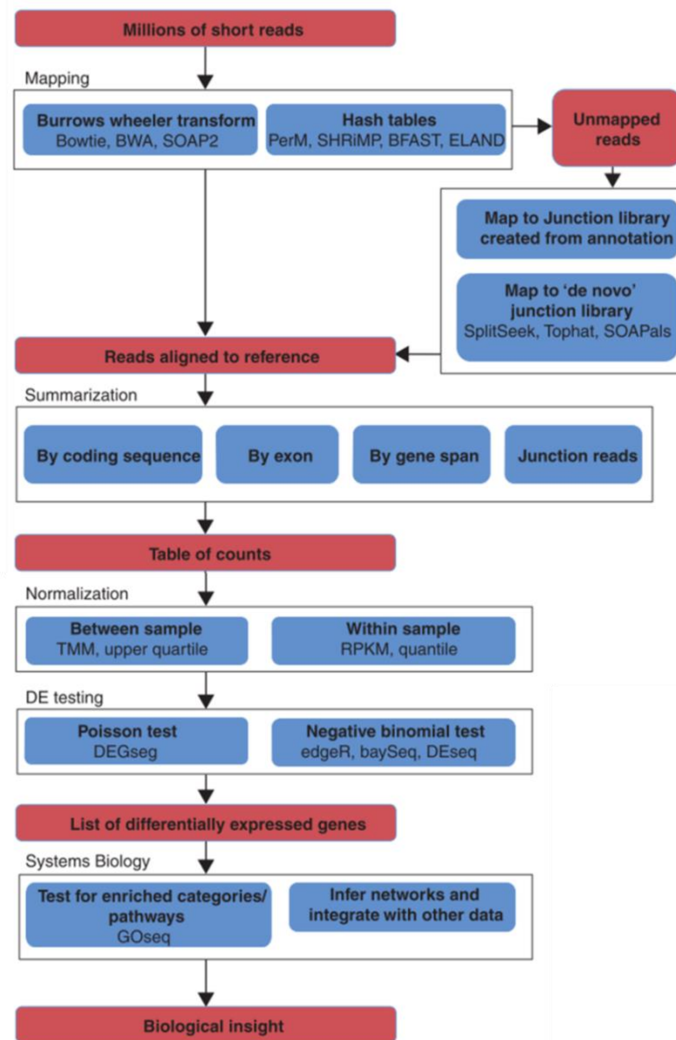


Figure 1.5. Overview of the RNA-seq analysis pipeline for detecting differential expression. The steps in the pipeline are in red boxes; the methodological components of the pipeline are shown in blue boxes and bold text; some software examples and methods for each step are shown by regular text in blue boxes (Extracted from Finotello F *et al*., 2014).

As a digital measure (count data), RNA-seq scales linearly even at extreme values and allows to capture a wider range of expression values providing also, information on RNA splice events (Mortazavi et al. 2008). This technique provides nearly unlimited possibilities in modern bioanalysis.

1.4.2. GENE EXPRESSION ESTIMATION

Quantification of gene expression can be performed by a wide range of methodologies. The first technique applied to measure the expression of a gene was the Northern blot hybridization (Streit et al. 2009). This method, assess the relative expression level of transcripts based on the intensity of the hybridization band. A more specific and accurate technique was developed later on, the reverse transcriptase quantitative PCR (RT-qPCR) (Tse & Capeau). This method allows performing either a relative or an absolute quantification of the expression level of a particular mRNA. Recently, several techniques such as expressed sequence tags (ESTs), serial analysis of gene expression (SAGE), differential display (DD), expression microarrays and high throughput sequencing (HTS) have emerged allowing researchers to analyze gene expression at a whole-genome level. There are two main quantification strategies: relative to an external standard curve or to one or more co-amplified internal control mRNAs (Pfaffl 2012).

- **Quantification relative to external standard curves**

This method is based on the use of a dilution series of an external standard, which is then used to generate a standard curve of threshold cycle (Ct) against an initial target copy number. Copy numbers of unknown samples can be estimated from the linear regression of that standard curve, considering the y-intercept giving the sensitivity and the slope giving the amplification efficiency (Bustin et al. 2005). To construct these standard curves, we can use from PCR fragments, single-stranded sense-strand synthetic oligodeoxyribonucleotides to commercially available universal reference RNAs and although they are often highly reproducible, external standards cannot detect or correct for inhibitors that may be present in the samples (Livak & Schmittgen 2001).

- **Quantification relative to internal standards or control mRNAs**

This method is based on the comparison of the Ct values from target RNAs to those of one or more internal reference genes. We obtain a ratio of the target-specific signal to the internal reference as a measure of its expression, which represents a corrected relative value that can be compared between samples. In this case, the amplification efficiencies of target and reference genes have to be similar, since this directly affects the accuracy of any calculated expression result. Several models have been published to correct for efficiency however, within relative quantification, the $2^{-\Delta\Delta CT}$ Ct method is the most used (Livak & Schmittgen 2001). This method assumes that the RT-qPCR efficiency is 100 % and is based on the use of reference genes whose expression is stable between the analyzed samples to correct for any difference in sample managing. Normalized expressions are then made relative to a calibrator sample which is randomly chosen. Since the expression of most reference genes vary significantly with treatment or between individuals, relative quantification can be confusing sometimes (Anon 2010).

1.4.2.1. Gene expression quantification by RT-qPCR

Reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR) is a powerful, sensitive and reliable technique that has become the gold standard for measuring and evaluating the expression of single or multiple genes. RT-qPCR is used when the starting material is RNA, either mRNA or other RNA source, which is first transcribed into complementary DNA (cDNA) by a reverse transcription reaction and then is used as the template for PCR amplification (Tse & Capeau). Ideally, PCR primers should be designed to span an exon-exon junction. This design reduces the risk of false positives from amplification of any contaminating genomic DNA, since the intron-containing genomic DNA sequence would not be amplified. If primers cannot be designed to separate exons or exon-exon boundaries, it is necessary to treat the RNA sample with RNase-free DNase I or dsDNase in order to remove contaminating genomic DNA (Bustin 2004). Once the RNAs are reverse transcribed to cDNAs, they are amplified by a variant of the polymerase chain reaction (PCR) used to simultaneously amplify and quantify the amplification product (Mullis et al. 1986). This quantification of the amplified product can either be absolute (number of molecules of a transcript) or relative

(proportion of transcripts in relation to one or more control genes) and it is based on the fluorescence emitted by an excited fluorochrome during the exponential phase of the reaction (Figure 1.6).

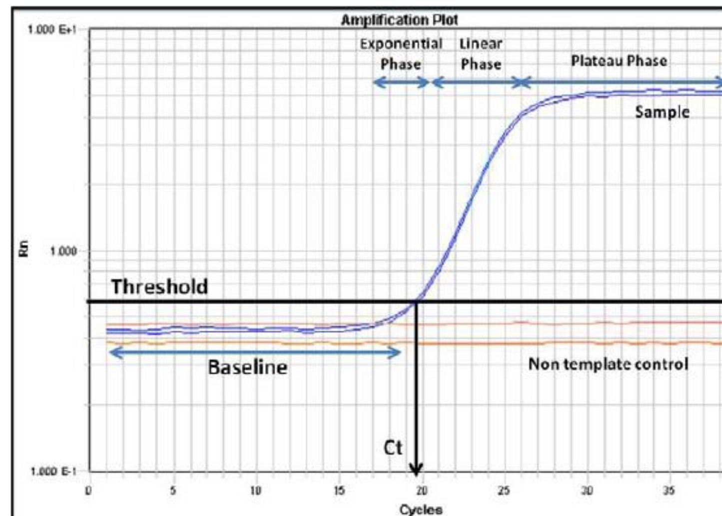


Figure 1.6. RT-qPCR amplification plot.

When the reaction reaches the exponential phase, the real-time PCR instrument estimates two values: the “threshold line”, which is the level of detection at which a reaction reaches a fluorescent intensity above background and the “cycle threshold” (Ct), which is the PCR cycle at which each sample reaches this level. This Ct is inversely proportional to the expression level. Low expression levels result in high Ct and the opposite (Bustin et al. 2005).

RT-qPCR techniques can be classified depending on the detection chemistries, which can be either specific (probe-based) or non-specific (non-probe based):

- **Non-probe based chemistry**

Techniques based on unspecific fluorochromes are based on the exponential detection of the produced double stranded DNA (dsDNA) by a fluorochrome that binds in a non-specific way to each double chain produced during the PCR. The most widely non-probe-based chemistry detects the binding of SYBR Green to dsDNA (Bustin 2000). In solution, this intercalating dye exhibits little fluorescence. However, when it binds to the dsDNA emits a strong fluorescent signal. The intensity of the fluorescence increases as the PCR products accumulate. This technique is the most economical and easiest to use. It allows assessing specifically amplified DNA fragments from the melting temperature (Tm) by analyzing the melting curves. However,

since the dye does not discriminate the double-stranded DNA from the PCR products and those from the primer-dimers, overestimation of the target concentration can be a problem (Schmittgen et al. 2000; Rajeevan et al. 2001)

- **Probe-based chemistry**

Techniques based on probe sequences that fluoresce upon hydrolysis or hybridisation use at least one fluorescently labeled primer (Juskowiak 2010). This probe is usually attached to a fluorochrome and a quencher and it hybridizes in the intermediate zone between the forward and the reverse primer; that is, within the amplicon. Thus, when the probe is intact, the proximity between the fluorochrome and the quencher inhibits the emission of fluorescence. In the other hand, when the fluorophore and the quencher are distant due to degradation of the probe by the 5'-3 'exonuclease activity of DNA polymerase, fluorescence is emitted (Schmittgen et al. 2000). This allows monitoring a change in fluorescence pattern only if the DNA sequence complementary to the specific probe is amplified. With this approach, any possible nonspecific amplification is eliminated. Most widely used probes are TaqMan® from Life Technologies and HybProbes from Roche.
To provide meaningful and reproducible results, parameters such as RNA extraction, RNA integrity, cDNA synthesis, primer design, amplicon detection, and data normalization have to be taken into account (Bustin et al. 2009).

1.4.2.2. MicroRNA expression quantification by RT-qPCR

Despite the small size of miRNAs, there are some technologies that enable high-throughput and sensitive miRNA profiling such as microarrays (Pradervand et al. 2009), real-time quantitative PCR (RT-qPCR) (Mei et al. 2012) and bead-based flow cytometry (Jang et al. 2011). Because of its accuracy and specificity, RT-qPCR has become the method of choice not only for measuring gene expression levels, but for analyzing the expression level of non-coding RNAs including miRNAs (Schwarzenbach et al. 2015). RT-qPCR for measuring miRNAs expression is based on the same concepts applied to determine mRNA expression (Mei et al. 2012). The challenge in adapting this technique to miRNA expression quantification resides on their short length, because miRNAs have approximately the same size as conventional RT-

qPCR primers (20-24 nt). To overcome this issue a subset of RT-qPCR methodologies have been developed: poly(A) miRNA-based RT-qPCR (Shi & Chiang 2005), stem-loop RT-qPCR (Chen et al. 2005; Mestdagh et al. 2008), Universal RT microRNA PCR (Ingrid Balcells, Cirera, et al. 2011) and miR-ID (Kumar et al. 2011) (Figure 1.7). These methodologies differ in some steps such as the cDNA synthesis (by using stem-loop primers, linear miRNA-specific primers or by tailing RNAs with E. coli Poly (A) Polymerase or T4 RNA Ligase 1), the amplicon detection (using SYBR Green or TaqMan probes) and the primer design, which is linked to both, the type of cDNA synthesized and the used method to detect the amplicon (Figure 1.8).

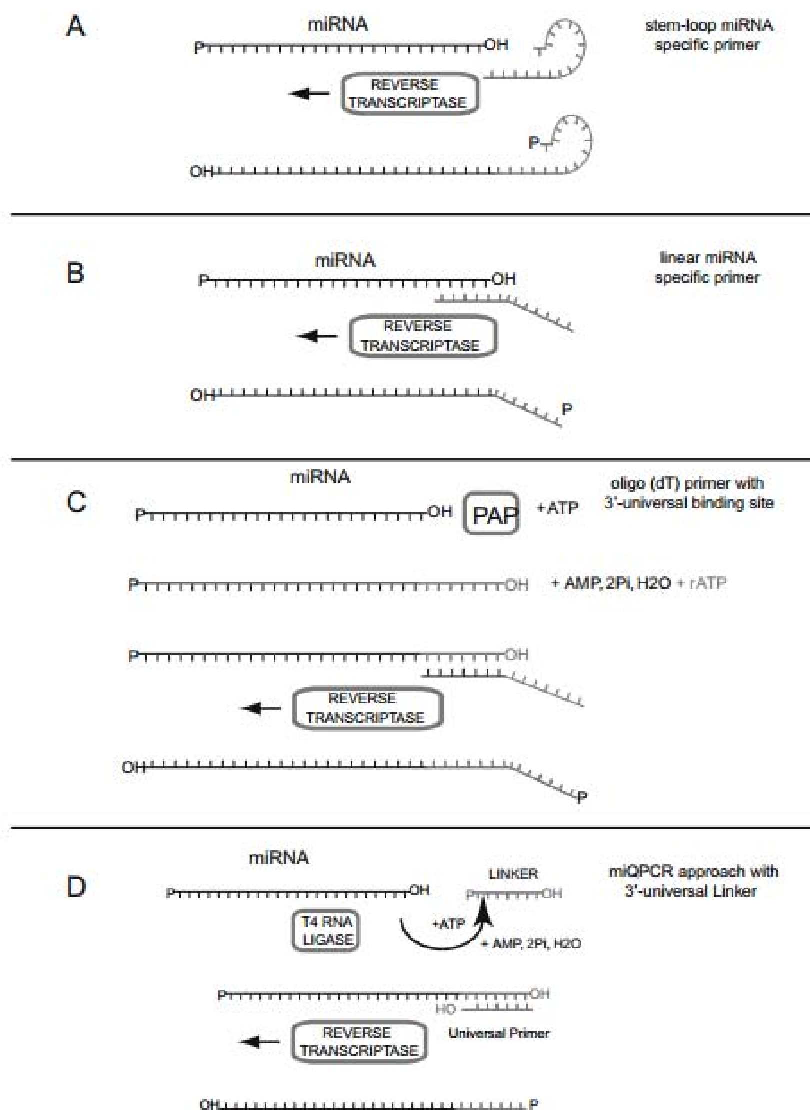


Figure 1.7. Alternative RT-qPCR methodologies to generate cDNA using stem-loop primers (A), linear miRNA-specific primers (B) or by enzymatic tailing using Poly(A) Polymerase (C) or T4 RNA Ligase (D). Modified from Benes *et al.*, 2010.

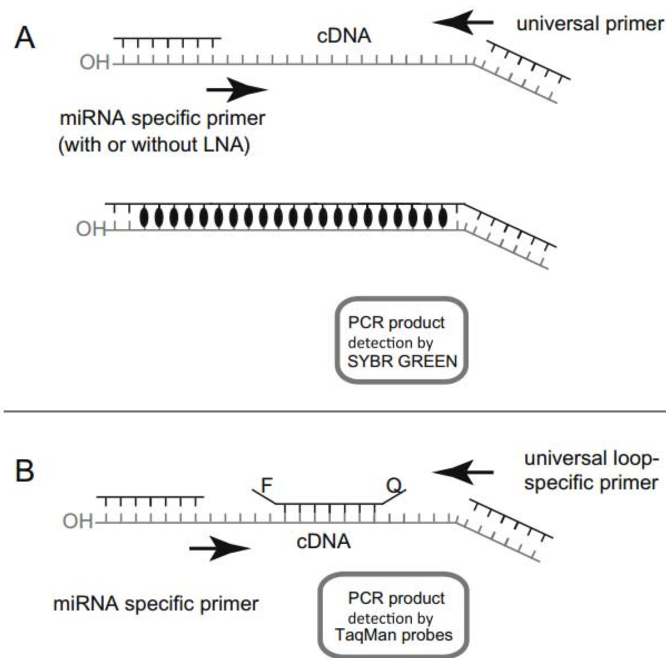


Figure 1.8. Amplicon detection by using SYBR Green (A) or TaqMan probes (B). Modified from Benes *et al*., 2010.

1.5. REGULATION OF GENE EXPRESSION

It is well known that gene expression in cells and tissues is not constant and requires precise spatial-temporal remodeling. Thus, to ensure an optimal response of the cell to its environment and the demands of the whole organism, gene expression regulation is completely necessary (Mata et al. 2005).

This regulation is controlled by several mechanisms acting mainly at two different levels: transcriptional and post-transcriptional. These steps are connected and coordinated, controlling gene expression from the initiation of transcription to protein translation (Dahan et al. 2011).

Traditionally, studies have been focused in the regulation at the transcriptional level because it was considered the most important step of gene expression and it was easy to study using the established methods (Mata et al. 2005). Regulation at this level is controlled by proteins that can be classified in two groups: sequence-specific DNA binding proteins, such as, transcription factors, and proteins of large multi-protein RNA polymerase machines, such as, TATA-binding proteins (Levine & Tjian 2003). However, there are other regulatory mechanisms that do not involve any change in DNA sequence which are called epigenetic mechanisms such as, DNA

methylation and histone modifications, which play an essential role as regulators of transcription (Bell & Spector 2011). More recently, it has been observed that post-transcriptional regulation provides a more rapid response to cellular signals and/or environmental stimulus than transcriptional regulation (López-Maury et al. 2008). Since its role in many biological processes and relevant diseases has been demonstrated, the importance of this mechanism has emerged (Y. Huang et al. 2011). However, post-transcriptional regulators are not completely understood. It is known that RNA binding proteins (RBPs) and non-coding RNAs (ncRNAs) are the main post-transcriptional mechanisms. Although there are several ncRNAs species, such as piwi-protein-interacting RNAs (piRNAs), endogenous short interfering RNAs (endo-siRNAs) and long noncoding RNAs (lncRNAs), microRNAs (miRNAs) are clearly the most important post-transcriptional regulators of gene expression (Strachan & Read 2011).

1.5.1. microRNAs

MicroRNAs (miRNAs) are a class of small non-protein coding RNAs of approximately 20–25 nucleotides (nt) long that act mainly as post-transcriptional down-regulators of protein-coding transcripts (Bartel 2004). These small RNAs comprise one of the more abundant classes of gene regulatory molecules. MiRNAs biogenesis is temporal and spatial dependent. They are transcribed by RNA polymerase II as parts of longer primary transcripts called pri-miRNAs, which are processed to mature miRNAs in two consecutive maturation steps. First, the 5' cap and 3' poly-A-tail of pri-miRNAs sequence is recognized by a multiprotein complex (microprocessor complex) formed by RNase III enzyme Drosha and the Di George Syndrome critical region gene 8 (DGCR8). This interaction ends up with the formation of a hairpin-structured RNA molecule of 70–100 bp called miRNA precursor or pre-miRNA (Lee et al. 2003). Then, the pre-miRNA is transported out of the nucleus by Exportin-5. Once at the cytoplasm, these double stranded miRNA precursors are processed by DICER giving rise a final 18-25 nt double-stranded RNA duplex which contains the mature miRNA guide strand and the passenger or miRNA* strand (Lee 2002; Lee et al. 2003). Finally, the mature miRNA guide strand is incorporated into a miRNA-protein complex, where it interacts with a member of the Argonaute (Ago) protein family forming the miRNA-induced silencing complex (miRISC or miRNP) (Figure 1.9). Despite advances in the understanding of the mechanisms operating

during miRNAs biogenesis, little is known about their stability and permanence. Most of them have a half-life of over 14 hours (Lee 2002); however, some miRNAs appear to have really fast kinetics of degradation suggesting a specific regulation mechanism for each miRNA or group of them (Rüegger & Großhans 2012; Bail et al. 2010).

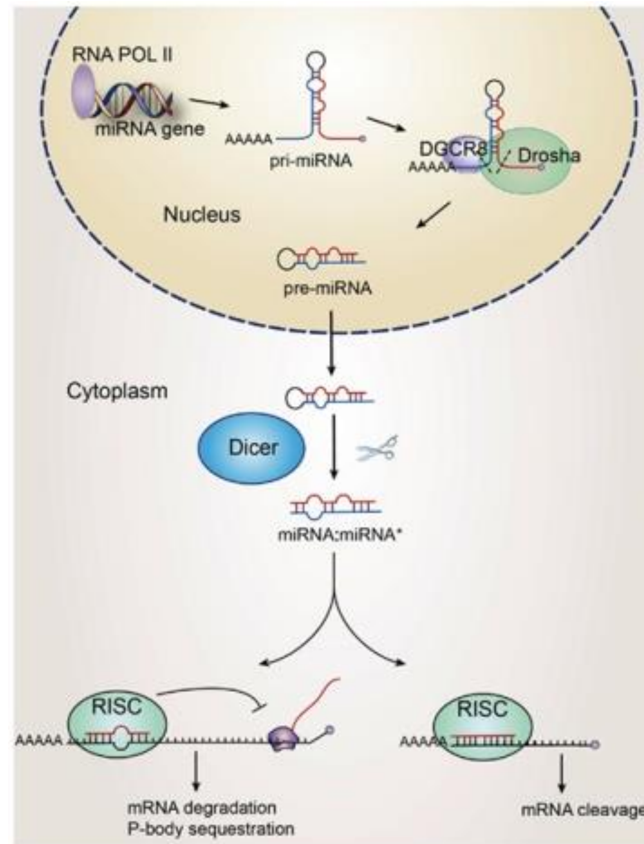


Figure 1.9. Biosynthesis of miRNAs (from Barca-Mayo et al. 2012).

1.5.1.1. miRNA targeting

Since the first mRNA regulated by a miRNA was discovered in the early 90s (Wightman et al. 1993), biochemical assays, genetic and bioinformatics have revealed many regions within the sequence of the mRNAs, susceptible to bind by base complementarity to miRNAs (Ekimler & Sahin 2014). When the mature miRNA guide strand forms the miRISC or miRNP guides this large protein complex to partial complementary target sites, which are typically located at the 3' untranslated region (UTR) of the target mRNA. The nucleotide sequence of the miRNA that specifically binds to the mRNA target site is located between positions 2 and 7 in direction 5'-3' and constitutes the "seed" region (Cai et al. 2009) (Figure 1.10). Those members of the same miRNA family present a high degree of sequence homology in this region and miRNA binding

sites are widely conserved in different species (Friedman et al. 2009). Although miRNA binding sites are mostly located in the 3 'UTR (Bartel 2004), less common functional sites in the 5' UTR have recently been reported, too (Brümmer & Hausser 2014).

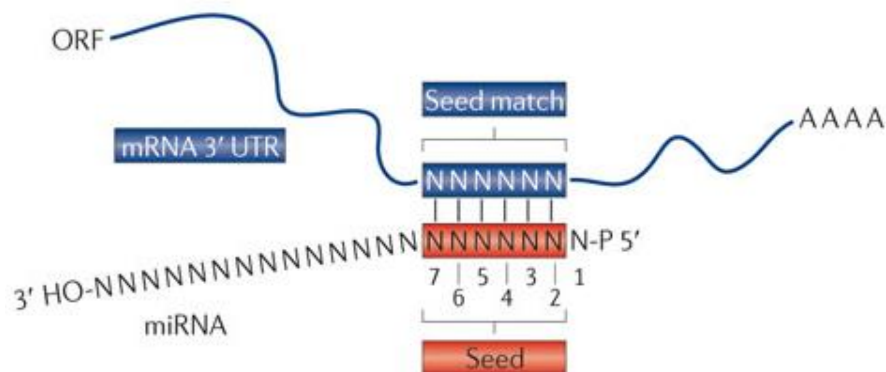


Figure 1.10. Target recognition by microRNAs in animals. Modified from Huntzinger *et al.*, 2011).

Complementarity of the "seed" region is essential for the repression of the target mRNA expression, however, additional interactions between flanking nucleotides of the "seed" region and the 3'UTR region of the mRNA, can increase specificity and stability (Brennecke et al. 2005). Other factors also influence the specificity of the miRNA:mRNA binding: Increasing AU bases near the "seed" region, additional pairing of nucleotides in the 3'UTR region of the mRNA, the proximity of several miRNAs binding sites within the 3' UTR of the target mRNA and the position of the miRNA binding site relative to the center of the UTR region and the stop codon (Grimson et al. 2007). The interaction between the mRISC complex and the target mRNA leads to the degradation or translational repression of the gene (Koscianska et al. 2011). The mechanisms by which miRNAs exert this repression have been in constant debate as they can vary depending on the experimental model used (Grimson et al. 2007). Besides acting as inhibitors of translation, some studies have also related increases in the amount of miRNAs with decreased populations of mRNAs in different organisms (Bagga et al. 2005; Valencia-Sanchez et al. 2006).

1.5.1.2. Functional validation of miRNA targeting

The commonly used methods to experimentally validate miRNA:mRNA interactions are the well-established techniques of qRT–PCR, luciferase reporter assays and western blot (Huang et al. 2010). Although the existence of multiple methodologies to establish miRNA targeting, the most extensively used are reporter assays. Genetic reporters are used as indicators in the study of gene expression and other cellular events related to gene expression (Juskowiak 2010). Normally, a reporter gene is cloned with a DNA sequence of interest into an expression vector that is then transferred into cells. After transfection, the presence of the reporter in the cells is analyzed by directly measuring the reporter protein itself or its enzymatic activity (Gould & Subramani 1988).

Most commonly used reporter genes are the luciferases from Firefly (*Photinus pyralis*) and Renilla (*Renilla reniformis*), because the luminescence of their resulting proteins can be easily detected (Jin et al. 2013). This bioluminescence is based on the interaction of the enzyme luciferase, with a luminescent substrate called luciferin in a chemical reaction that takes place in two steps:

Step 1: luciferin + ATP → luciferyl adenylate + PPi

Step 2: luciferyl adenylate + O_2 → oxyluciferin + AMP + light

Light is emitted because the reaction forms oxyluciferin in an electronically excited state. The reaction releases a photon of light as oxyluciferin returns to the ground state allowing a quantitative measure of its expression (Thorne et al. 2010). As biological samples are intrinsically complex, data available from a single reporter may be insufficient for achieving reliable results. For this reason, we have performed our experiments using a Dual-Luciferase® Reporter Assay system (Promega) which provides more efficient results when testing the interaction of microRNA and mRNA. This dual system enables the sequential measurement of both firefly and Renilla luciferases from each sample (Thorne et al. 2010), which makes possible to differentiate genetic responses of interest from non-relevant influences in the experimental system (Figure 1.11).

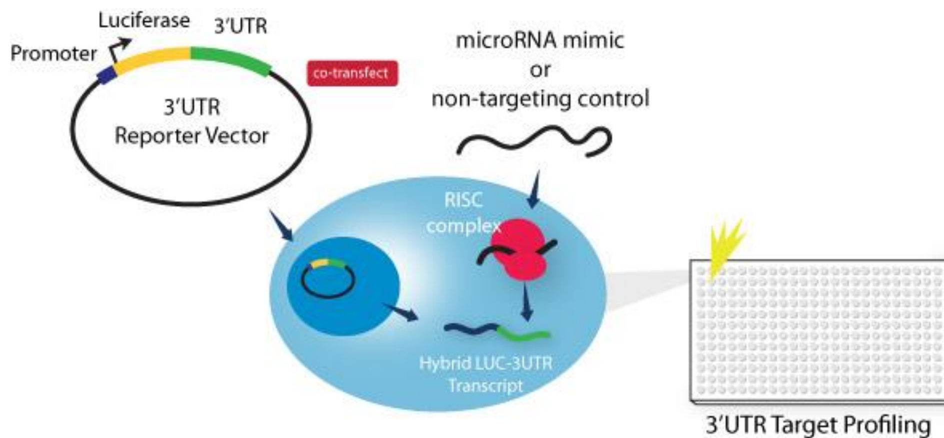


Figure 1.11. Mechanism to detect the impact of a miRNA on the regulation of a target mRNA by luciferase reporter assay (from: http://www.biocat.com/reporter-assays#3UTR_Constructs).

1.5.2. Regulation of miRNA processing

MicroRNAs processing and maturation can be directly affected by the presence of genetic variants in miRNA genes leading to a dysregulation of their expression level. Recently, it has been demonstrated that single nucleotide polymorphisms (SNPs) in miRNA genes (miR-SNPs) can alter their function by modulating one or more of their processing steps (Mishra et al. 2008; Ryan et al. 2010).

Several authors have directly associated miR-SNPs with many relevant diseases (Zhang et al. 2015; Li et al.; Smith et al.; Hu et al.; Yang et al. 2010; Z. Hu et al. 2008), suggesting them as putative biomarkers to predict disease risk and/or prognosis (Xu & Tang 2015; Z. Hu et al. 2008). Polymorphisms in either the primary or precursor form of a miRNA have relevant functional implications (Z. Hu et al. 2008) and could affect mature microRNA expression either positively or negatively (Han et al. 2013). Changes at the sequence level can affect both, the processing of the precursor molecules during mature miRNA biogenesis, and the interaction miRNA:mRNA by altering the recognition and binding process. Yang and collaborators demonstrated that editing pri-microRNA-142 seems to interfere further processing by Drosha / DGCR8 complex, causing a decrease in the expression of *miR-142-3p*, and *miR-142-5p* mature forms (Yang et al. 2006). Similar results were observed in the processing of *miR-151* precursor in brain (Kawahara et al. 2007). Moreover, recent in vitro tests have also shown that nucleotides of the "seed" themselves, might be able to repress translation of mRNAs (Obad et

al. 2011), demonstrating that variations within this region could affect the activity of miRNAs, altering the translational repression of a specific mRNA. Thus, the allelic variation of miRNA target sites and in miRNAs themselves is thought to be a contributing factor to many phenotypic differences observed in livestock (Liu et al. 2010).

1.5.3. The role of miRNAs in reproduction

Profiling studies in livestock have revealed that miRNAs have key functions in essential biological processes, including cellular differentiation, proliferation, and apoptosis (Kotlabova et al. 2011; Kloosterman & Plasterk 2006), which are relevant in embryo formation, early development, and implantation (Viganò et al. 2003). Although the exact role of miRNAs in normal embryo formation and endometrial preparation for pregnancy still remains unclear, they have been widely associated with mammalian development (Tang et al. 2007). Moreover, Yu *et al.*, demonstrated that miRNA expression in mouse embryos was higher than in mature mouse tissues, confirming their role during embryo development (Yu et al. 2007).

To date, only a few reports have explored miRNAs expression profiles in porcine reproductive tissues, and despite miRNAs function has been related to endometrial receptivity (Sha et al. 2011; Altmäe et al. 2013; Xia et al. 2014), implantation (Chakrabarty et al. 2007; S.-J. Hu et al. 2008; Revel et al. 2011; Su et al. 2014), labor and spontaneous fetal loss in pigs (Montenegro et al. 2009; Williams, Renthal, Condon, et al. 2012; Williams, Renthal, Gerard, et al. 2012; Renthal et al. 2010; Hassan et al. 2010), miRNA-mediated regulation of sow's pregnancy remains unclear.

1.6. MEIBMAP INTERCROSS

Large genetic variation has been found among porcine breeds regarding litter size (Bradford 1979). The most extreme phenotypes have been observed between European and Asian breeds, which differ significantly in their prolificacy levels. The Meishan breed is considered one of the most prolific porcine breeds with an average of 14.3 piglets born alive per parity (Bidanel 1993), whereas the Iberian breed is considered a very low-prolificacy breed with an average of 7 piglets per parity (Silió L. 2001).

Within the MEIBMAP project, an experimental F_2 intercross was created by mating 18 Meishan (Domaine du Magneraud, INRA, France) sows and three Iberian boars from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain), which generated an F_1 composed of eight boars and 97 sows that were intercrossed to obtain the F_2 progeny (Figure 1.12). Among the 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. Over four consecutive parities, the total number of piglets born (TNB) and the number of piglets born alive (NBA) were obtained for each sow, with an average 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.

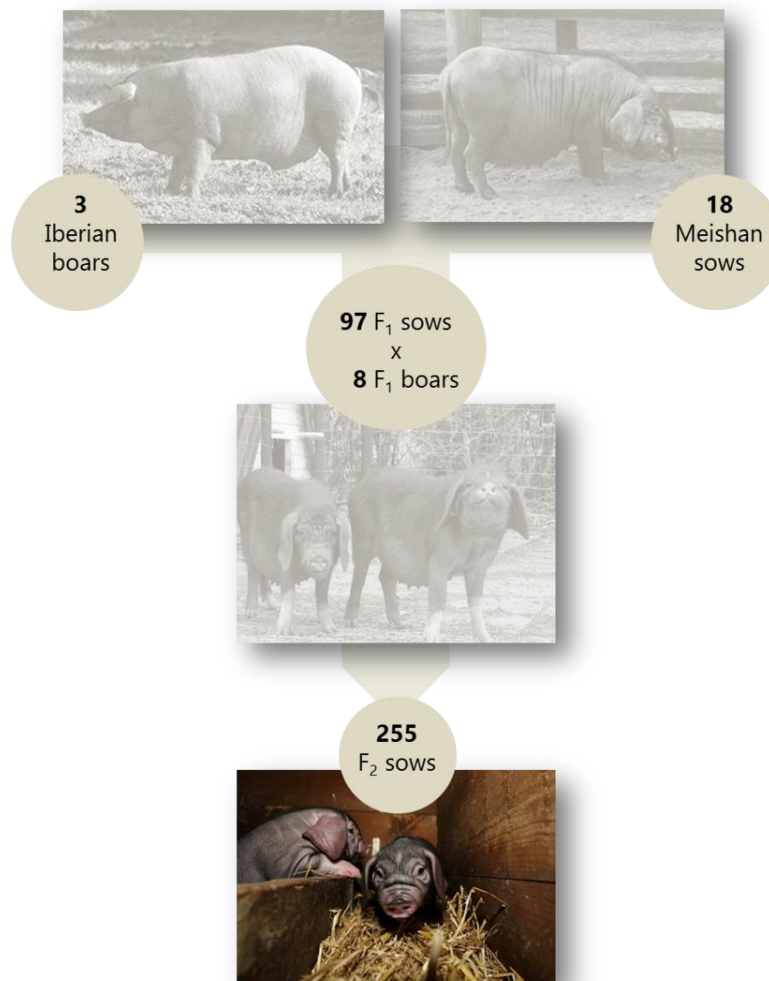


Figure 1.12. Experimental Iberian x Meishan intercross (MEIBMAP).

1.6.1. Reproductive QTLs previously identified in the MEIBMAP project

Within the MEIBMAP project several QTLs related with reproductive traits were identified (Table 1.5). Three significant QTL associated with the teat number trait were identified in SSC5, SSC10 and SSC12 (Rodriguez et al. 2005), which accounted for the 30% of the total phenotypical variance for this trait. At the SSC8, a QTL with a significant effect on the gestation length in sows was also identified, and it was also demonstrated that the substitution of the Iberian allele increased gestation length in 0.5 days (Casellas et al. 2008).

Afterwards, two highly significant QTL in SSC13 and SSC17 were identified at the genome wide level for NBA and TNB traits (Noguera et al. 2009). These two QTLs presented high additive and explain around 2 % to of the phenotypic variance. The discovery of 18 epistatic interactions for NBA and TNB involving 13 out of the 18 pig autosomes (Noguera et al. 2009) demonstrates that the phenotypic variance of these particular traits can be highly influenced by a complex network of interacting loci. One year later, Fernández-Rodríguez validated the two epistatic QTL interactions on SSC12 for NBA and TNB traits (Fernández-Rodríguez et al. 2010).

Table 1.5. Significant QTL for reproduction traits analyzed within the MEIBMAP project.

Trait[a]	SSC[b]	Position cM	Reference
TN	5	29	Rodríguez *et al.*, 2005
	10	71	
	12	70	
GL	8	110	Casellas *et al.*, 2008
NBA	13	50	Noguera *et al.*, 2009
	17	22	
TNB	13	55	
	17	22	
NBA	12	15	Fernández-Rodríguez *et al.*,2010
		91	
TNB	12	14	
		91	

[a]TN = Teat Number; GL = Gestation Length; NBA = Number of piglets born alive; TNB = Total number of piglets born; [b]*Sus Scrofa* chromosome.

1.6.2. Candidate genes previously identified in the MEIBMAP project

During the MEIBMAP project, several approaches have been used to identify those candidate genes involved in the regulation of prolificacy related traits (Table 1.6). In the first studies carried out by Tomás and collaborators and later on by Ramírez et al., three main candidate genes associated with new-born piglet vitality during the first hours after birth were identified: the prolactin receptor (*PRLR*), the dopamine β-hydroxylase (*DBH*) and the vascular cell adhesion molecule 1 (*VCAM1*) (Tomas 2006; Tomás, Casellas, et al. 2006; Ramirez et al. 2008).

For the *PRLR* gene, the obtained results also indicated an effect of this gene on the ovulation rate (Tomas 2006). None of these genes could be associated with litter size traits, however they found this associations for four genes: the melatonin receptor 1A (*MTNR1A*), located within a QTL for litter size traits on SSC17 and associated with litter size traits with additive and dominant effects that change depending on the season of the year, for bone morphogenetic protein receptor type 1β (*BMPR1B*), for which only suggestively effects on litter size during the first parity were determined (Tomás, Frigo, et al. 2006) and finally, the solute carrier family 9 (sodium/hydrogen exchanger) member 3 regulator 1 (*SLC9A3R1*) and the inducible nitric oxide synthase 2 (*NOS2*), both located within the confidence interval of the two epistatic QTL affecting litter size detected on *SSC12* (Fernández-Rodríguez et al. 2010). Moreover, although in the sequence of the *SLC9A3R1* the authors identified a polymorphism that was discarded to be the causal mutation, for the polymorphism identified in the sequence of *NOS2* gene, the results suggested that the *NOS2* haplotype could be the causal mutation underlying QTL2 on SSC12.

Later on, Martínez and collaborators based their studies in the parathyroid hormone-like hormone (*PTHLH*) gene as a candidate due to its location on the SSC5 where a QTL for teat number was already identified (Martínez-Giner et al. 2011). Despite this, no association could be detected between this gene and the teat number phenotype in the studied sows.

In 2011, Balcells et al. selected as candidate genes three porcine genes of the same family: *ITIH-1*, *ITIH-3* and *ITIH-4* (I Balcells et al. 2011). These genes have a reported function in many reproductive processes and have a positive mapping into the confidence interval of the QTL associated with litter size described in SSC13 (Noguera et al. 2009). Their results led to the identification of ten SNPs in the ITIH cluster that were proposed as potential markers to be used for selection for litter size in pigs. In this study it was also observed that higher expression levels of *ITIH-3* were specifically related to prolificacy levels in the MEIBMAP F_2 sows.

In the same context, the porcine *MUC4* gene located on SSC13 was analyzed. In humans, its expression has been associated with endometriosis and infertility (Chang et al. 2011) and in pigs, it has been proposed as a potential regulator of placentation (Østrup et al. 2010; Govindarajan & Gipson 2010). Balcells and collaborators identified a *MUC4* polymorphism (DQ124298:g.344A>G) that was associated with litter size (Ingrid Balcells, Castelló, et al. 2011). Moreover, their results showed a differential expression of this gene regarding the number of embryos attached to the uterus at day 30 of gestation, suggesting that it may participate in the establishment of an optimal uterine environment essential for successful embryo development during the early stages of gestation.

Table 1.6. Candidate genes analyzed within the MEIBMAP project. Location in *Sus Scrofa* Chromosome (SSC), QTL associated with prolificacy related traits identified by mapping the candidate gene position (QTL) and polymorphism changes at nucleotide and aminoacidic level (in brackets).

Gene	SSC	QTL	SNP reference	Traits	Reference
BMPR1β	8	-	G804C C852T C960T	NBA (1p)† NW (1p)†	Tomás *et al.*, 2006c
DBH	1	-	A463G(Thr155Ala) A510C T612C A616G(Lys206Glu) C744T	RT-1* TS† WB†	Tomás *et al.*, 2006b
ESR1	1	-	PvuII[1]	NBA* TNB*	Braglia *et al.*, 2006
ITIH-1,-3,-4	13	TNB/NBA	-	NBA* TNB*	Balcells *et al.*, 2011
MTNR1A	17	TNB/NBA	T162C	TNB* NBA*	Ramírez *et al.*, 2009
MUC-4	13	TNB/NBA	DQ124298:g.344A>G	NBA* TNB*	Balcells *et al.*, 2011
NOS2	12	TNB/NBA	A662G A1791C C2192T	NBA* TNB*	Fernández-Rodríguez *et al.*, 2010
PRLR	16	-	C1217T(Leu406Pro) C1283A(Asp428Ala) G1439A(Lys480Arg) T1528A(Met510Leu) G1600A(Gly534Ser) G1789A(Gly597Ser)	CL* RT-1* HR-0† OS-1† TS†	Tomás *et al.*, 2006a
PTHLH	5	TN	C56T(Ser19Leu)	TN	Martínez-Giner *et al.*, 2011
SLC9A3R1	12	TNB/NBA	A839G A259	NBA* TNB*	Fernández-Rodríguez *et al.*, 2010
VCAM1	4	-	T306A(Asn102Lys) C558T	TS* TU†	Ramírez *et al.*, 2008

TN = Teat number; TNB = total number piglets born; NBA = number of piglets born alive. 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 thfier st 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.[1]Polymorphism described by Rothschild *et al.,* (1996).*Significant (p-value < 0.05); †suggestive (p-value < 0.1).

2. OBJECTIVES

This thesis has been performed in the context of the MEIBMAP project (AGL2010-22358-C02-01/AGL2004–08368-C03) founded by the Ministry of Economy and Competitiveness and the Consolider-Ingenio 2010 Programme (CSD2007-00036), both form the Spanish Government.

The present research is the result of a coordinated project involving IRTA, INIA and UAB whose main goal has been the study of the genetic and molecular underpinnings that contribute to reproductive efficiency in swine.

The specific objectives of this thesis were:

1. **Identify key differences in gene expression associated to swine reproductive efficiency in the endometrium of pregnant sows at day 30 - 32 of its gestation.**

1.1. Determine if the differentially expressed genes are directly involved in relevant pathways for pregnancy establishment and successful embryo development.

2. **Explore the regulatory mechanisms that mediate the expression of reproductive-related genes.**

2.1. Identify key differences in miRNA expression associated to extreme prolificacy levels.

2.2. Determine if the presence of single nucleotide polymorphisms in the sequence of reproduction-related precursor miRNAs causes expression differences of the mature miRNA.

2.3. Estimate the association of miR-SNPs with the observed reproductive efficiencies between the Iberian and Meishan pig breeds.

2.4. Genotype the whole F_2 population for the identified variants and conclude if the genotype is a determinant factor for the sow EBV.

3. **Functionally validate if miRNA:mRNA interactions constitute the major mechanism of gene regulation.**

3.1. Confirm the interaction between relevant reproduction-related genes and their target miRNAs.

3.2. Confirm if the interaction with a target miRNA causes the down-regulation of candidate reproduction-related genes.

3.3. Establish the association between this down-regulation and the observed differences in the reproductive efficiency of Iberian and Meishan pig breeds.

3. ARTICLES AND STUDIES

3.1. ENDOMETRIAL TRANSCRIPTOME PROFILING

3.1.1. Analysis of gene expression differences between extreme prolificacy phenotypes

Endometrial gene expression profile from pregnant sows with extreme phenotypes for reproductive efficiency

S. Córdoba, I. Balcells, A. Castelló, C. Ovilo, JL. Noguera, O. Timoneda, A. Sánchez

Scientific Reports, 2015 – Vol 5, p.14416

SCIENTIFIC REPORTS

OPEN

Endometrial gene expression profile of pregnant sows with extreme phenotypes for reproductive efficiency

Received: 27 April 2015
Accepted: 19 August 2015
Published: 05 October 2015

S. Córdoba[1], I. Balcells[1], A. Castelló[1], C. Ovilo[2], J. L. Noguera[3], O. Timoneda[1] & A. Sánchez[1]

Prolificacy can directly impact porcine profitability, but large genetic variation and low heritability have been found regarding litter size among porcine breeds. To identify key differences in gene expression associated to swine reproductive efficiency, we performed a transcriptome analysis of sows' endometrium from an Iberian x Meishan F_2 population at day 30–32 of gestation, classified according to their estimated breeding value (EBV) as high (H, EBV > 0) and low (L, EBV < 0) prolificacy phenotypes. For each sample, mRNA and small RNA libraries were RNA-sequenced, identifying 141 genes and 10 miRNAs differentially expressed between H and L groups. We selected four miRNAs based on their role in reproduction, and five genes displaying the highest differences and a positive mapping into known reproductive QTLs for RT-qPCR validation on the whole extreme population. Significant differences were validated for genes: *PTGS2* ($p = 0.03$; H/L ratio = 3.50), *PTHLH* ($p = 0.03$; H/L ratio = 3.69), *MMP8* ($p = 0.01$; H/L ratio = 4.41) and *SCNN1G* ($p = 0.04$; H/L ratio = 3.42). Although selected miRNAs showed similar expression levels between H and L groups, significant correlation was found between the expression level of *ssc-miR-133a* ($p < 0.01$) and *ssc-miR-92a* ($p < 0.01$) and validated genes. These results provide a better understanding of the genetic architecture of prolificacy-related traits and embryo implantation failure in pigs.

Pig is economically one of the most important species. Reproductive traits such as fertility and prolificacy can directly impact porcine profitability, becoming one of the most relevant traits from a genetic and economic point of view. The annual production of a sow is determined to a large degree by its litter size in terms of total number of piglets born (TNB) and number of piglets born alive (NBA) per parity. Total number of piglets born and NBA are the most important reproductive traits used in swine breeding programmes[1].

Although sow's fertility depends directly on the ovulation rate (OR), litter size is not strongly determined by this factor, but by the capacity of maintaining viable embryos throughout gestation. Prenatal mortality could be a determinant factor for litter size in pigs[2,3]. The relevance and timing of embryonic and foetal losses during gestation have been reported in many studies, and it is estimated that about 25–45% of fertilized ova do not survive through gestation. Losses of embryos and foetuses occur at each stage of development and are primarily determined by the uterine capacity of the pregnant sows[4]. A large genetic variation has been found among porcine breeds regarding litter size, being the Chinese Meishan one of the most prolific pig breeds known[5].

[1]Departament de Genètica Animal, Centre de Recerca en Agrigenòmica (CRAG), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Spain. [2]Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (SGIT-INIA), 28040 Madrid, Spain. [3]Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), 25198 Lleida, Spain. Correspondence and requests for materials should be addressed to S.C. (email: sarai.cordoba@cragenomica.es)

Improvements in litter size across the swine industry have occurred through different selection schemes such as phenotypic, family index, best linear unbiased prediction or hyper-prolific-based selection methods[2]. Being a complex trait regulated by a large number of genes, along with its low heritability, has made the selection of this character rather challenging for a number of years[6]. To date, main used strategies to detect those genes affecting litter size and its components have been: linkage analyses based on the identification of genomic regions linked with a phenotypic reproduction trait and candidate gene approaches, based on a priori knowledge of a gene having a high probability to play a relevant role in reproduction by their physiological role or location[7].

Significant quantitative trait loci (QTL) associated with porcine reproductive traits have been identified in our study population and many others: SSC3, SSC8, SSC9, SSC10 and SSC15 for ovulation rates[8–11], SSC7, SSC8, SSC12, SSC13, SSC14 and SSC17 for total number piglets born[6,12,13], SSC4 and SSC13 for number of stillborn[10,14] and SSC8 for uterine capacity and prenatal survival[13,15]. Although there are even more QTLs reported for litter size component traits, most of these results are inconsistent and true causal genes still remain scant due to the large disequilibrium linkage blocks present in the genome of livestock species[16].

In recent years, the knowledge obtained by deciphering the pig genome and advances in molecular genetics, such as the transcriptomic analysis by RNA sequencing, have provided a powerful tool to better understand the genetic architecture of prolificacy-related traits. Recent years have seen a remarkable rise in porcine transcriptomic data. The use of microarrays and large-scale transcriptome analysis to identify differentially expressed genes in specific tissues, cell types or breeds has shed light on many aspects of porcine production traits[17–24]. Despite this, there have only been a few comparative studies on uterine function for prolific pigs and a low number of experiments regarding differences in endometrial gene expression between porcine breeds[25–27].

In swine, during the oestrus cicle and throughout pregnancy many critical morphological and secretory changes take place in the uterus. These sets of physiological changes are clear evidence of the extremely complex interactions taking place between gene products and of remarkable transcriptomic reorganization. This highlights the importance of performing profiling experiments in porcine breeds with extreme prolificacy phenotypes, in order to better understand those gene interactions and the regulatory mechanisms affecting litter size in pigs.

An important mechanism of gene expression regulation is miRNAs. It is well known that miRNAs have key functions in many relevant biological processes, including cellular differentiation, proliferation, and apoptosis[28]. All these processes are involved in embryo formation, early development, and implantation. Although the exact role of miRNAs in normal embryo formation and endometrial preparation for pregnancy still remains unclear, they have been widely associated with mammalian development[29]. Moreover, Yu *et al.*, demonstrated that miRNA expression in mouse embryos was higher than in mature mouse tissues, confirming their role during embryo development[30].

The goal of our study is, then, to define those genes and miRNAs that are differentially expressed in the uterine endometrium of pregnant sows with extreme prolificacy phenotypes in an Iberian x Meishan F_2 population. These two porcine breeds differ significantly in their prolificacy levels, being the Meishan breed one of the most prolific porcine breeds, with an average of 14.3 piglets born alive per parity[31], whereas the Iberian breed is considered a very low-prolificacy breed with an average of 7 piglets per parity[32]. This makes our study population highly suitable for further investigating the biological underpinnings that contribute to controlling litter size in pigs.

Results

Differential gene expression. Uterine receptivity to implantation is a process that can be very different, depending on the species, but always involves several changes in the expression of genes that are directly involved in pathways, such as progesterone and oestrogen biosynthesis, immune recognition, membrane permeability, angiogenesis and vasculogenesis, transport of nutrients and signalling for pregnancy recognition. Thus, changes in the expression level of those genes may influence uterine receptivity to implantation. Analysis of read counts revealed a total of 141 differentially expressed genes (DEG) between high- and low-prolificacy samples when a false discovery rate (FDR) corrected *q-value* of 0.05 was set as the threshold for significance (supplementary table S1). Expression differences between H and L groups ranged from 5.61 to −5.84 fold. A total of 55 transcripts showed an overexpression in the high-prolificacy group, with expression differences ranging from −1.45 to −5.84 fold, whereas 49 showed an overexpression in the low-prolificacy group, with expression changes ranging from 1.51 to 5.61 fold. Moreover, we identified 27 transcripts expressed uniquely in the L group (2 annotated genes and 25 unannotated transcripts) and 10 transcripts expressed uniquely in the H group, including 4 annotated genes and 6 unannotated transcripts (See supplementary table S2).

Functional annotation and QTL mapping analysis. In order to establish whether differentially expressed genes found were involved in a relevant biological process for any stage of pregnancy establishment and development in the pig, we performed a gene ontology (GO) annotation and enrichment analysis. Obtained results revealed that the top over-represented functions were related with female pregnancy (*q-value* = 0.0001), maternal placenta development (*q-value* = 0.024) and decidualization (*q-value* = 0.024). All *p-values* were estimated through Chi square analysis and FDR corrected. An

GO term	Biological process	Log OR[a]	p-value	q-value[b]	DEG involved
GO:0007565	Female pregnancy	2.892	0.00000	0.0001	8
GO:0001893	Maternal placenta development	4.035	0.00004	0.0243	3
GO:0046697	Decidualization	4.218	0.00002	0.0243	3
GO:0048545	Response to steroid hormone	2.100	0.00004	0.0243	7
GO:0000038	Long-chain fatty acid metabolism	3.812	0.00007	0.0307	3
GO:0006694	Steroid biosynthetic process	2.267	0.00025	0.0722	5
GO:0009888	Tissue development	1.284	0.00030	0.0722	12
GO:0001503	Ossification	1.993	0.00026	0.0722	6
GO:0042127	Regulation of cell proliferation	1.264	0.00035	0.0722	12
GO:0060348	Bone development	1.957	0.00032	0.0722	6
GO:0009725	Response to hormone	1.605	0.00036	0.0722	8
GO:0001501	Skeletal system development	1.578	0.00043	0.0785	8
GO:0043129	Surfactant homeostasis	4.314	0.00056	0.0942	2
GO:0007398	Ectoderm development	1.813	0.00066	0.0969	6
GO:0051216	Cartilage development	2.396	0.00066	0.0969	4

Table 1. Functional enrichment analysis showing the top significantly-over-represented GO terms in which identified DEG are involved. [a]Odds ratio logarithmic transformation. [b]Benjamini-Hochberg FDR-corrected *p-value*.

FDR-corrected *q-value* of 0.10 was set as the threshold for significant functional enrichment (See table 1). We also performed this enrichment analysis considering separately those genes overexpressed in either group. The DEG overexpressed in H prolificacy samples were clustered in seven enriched general biological processes, including mainly: positive regulation of cell proliferation (GO: 0008284; *q-value*= 3.67E-06) and response to hypoxia (GO: 0001666; *q-value*= 0.0002). Differentially expressed genes showing an overexpression in L prolificacy samples were clustered in 11 enriched general biological processes, including mainly: proteolysis/cell-cell signalling (GO: 0006508, GO: 0007267; *q-value* = 2.36E-06) and *in utero* embryonic development (GO: 0001701; *q-value*= 0.0001).

In order to focus on those genes that could be strongly associated with reproduction and have an impact on litter size variation, a chromosomal localization of DEGs within known QTL intervals was performed. We identified a total of 59 mapping into known reproductive QTLs. Among them, 25 were located within a QTL specifically related with litter size: total number of piglets born alive (NBA), total number of piglets born (TNB), total number of piglets stillborn (TSB), body weight at birth (BW), body weight at 10 weeks (WT), body weight at weaning (WWT), mummified pigs (MMUM) and/or ovulation rate (OVRATE). Results are shown in supplementary table S3.

Candidate genes selection and expression levels validation: RT-qPCR. Among the 141 genes found differentially expressed in the RNAseq analysis (*q-value* < 0.05), we selected those displaying the most extreme differences between H and L groups (fold change ≥ 3) reducing the initial set to 28 genes. Based on the results obtained after the QTL mapping, we considered only those that have a positive mapping into known reproductive QTLs, reducing this number to 14 genes. Finally, considering the gene ontology (GO) annotation and enrichment analysis results and based on their known role in any relevant pathway related with reproduction, pregnancy or embryonic development, we chose 5 candidates: *HPGD*, *MMP8*, *PTGS2*, *PTHLH* and *SCNN1G* (See Table 2). Expression data obtained by RNA sequencing for these candidate genes was validated by RT-qPCR in 36 extreme individuals (H, n = 18; L, n = 18) of our F_2 population. We confirmed significant differences in the expression level of four of these five genes between H and L samples with an H/L ratio > 3.5: *MMP8* (mean H = 0.174, mean L = 0.035; *p-value* = 0.011), *PTGS2* (mean H = 0.144, mean L = 0.038; *p-value* = 0.026), *PTHLH* (mean H = 0.126, mean L = 0.033; *p-value* = 0.034) and *SCNN1G* (mean H = 0.117, mean L = 0.031; *p-value* = 0.048). Results are shown in Fig. 1a. The observed ratios between the expression level of selected candidate genes were similar in our RNAseq and RT-qPCR analysis: *HPGD* (RNAseq FC = 1.85, RT-qPCR FC = 1.81), *PTGS2* (RNAseq FC = 4.06, RT-qPCR FC = 3.79), *PTHLH* (RNAseq FC = 4.32, RT-qPCR FC = 3.78) and *SCNN1G* (RNAseq FC = 3.65, RT-qPCR FC = 3.72). Only for the *MMP8* gene the observed ratios between both analysis were slightly different (RNAseq FC = 2.99, RT-qPCR FC = 4.92).

Differential miRNA expression and *in silico* target prediction. The observed differences in the expression level of these genes between H and L prolificacy groups suggests that a different regulation mechanism may be occurring. We hypothesize that known gene regulators such as miRNAs could be

Gene	Function	References	QTL	Enriched Biological Process	RNA-seq analysis[a] High (RPKM)	RNA-seq analysis[a] Low (RPKM)	Log2FC	q-value	RT-qPCR analysis[b] High (RQ)	RT-qPCR analysis[b] Low (RQ)	FC	p-value
HPGD	Biosynthesis of prostaglandins (PTG)	Atli MO *et al.* 2010 Palliser HK *et al.* 2014 Kowalewski MP *et al.* 2014	–	Female pregnancy (GO:0007565)	102.14	28.40	1.85	0.032	0.271	0.160	1.69	0.118
MMP8	Collagen metabolism and preeclampsia Remodeling of the cervical and fetal membrane ECM	Mousa AA *et al.* 2012 Wang H *et al.* 2004	NNIP	Embryo development (GO:0009790)	68.49	8.64	2.99	0.008	0.152	0.034	4.44	0.013
PTGS2	Converts arachidonic acid to PGH2 Rate-limiting enzymes in PG synthesis Essential to reproduction	Waclawik A. *et al.* 2011 Blitek A. t al. 2006 Blitek A. *et al.* 2006 Sales KJ *et al.* 2003 Murakami M. *et al.* 2004 Lim H *et al.* 1997 Langenbach R *et al.* 1999 Silver RM *et al.* 1995	GEST	Maternal placenta development (GO:0001893)	109.91	6.57	4.06	0.008	0.129	0.037	3.50	0.027
PTHLH	Nipple development during pregnancy Preimplantation Fetoplacental development Embryonic mammary development	Martinez-Giner M *et al.* 2011 Guo L *et al.* 2012 Thota CS *et al.* 2005 Hiremath M *et al.* 2013	NSB	Lactation	286.09	14.36	4.32	0.008	0.108	0.029	3.69	0.027
SCNN1G	Preeclampsia	Marino G *et al.* 2013	OVRATE BW	Response to hypoxia (GO:0001666) Sodium ion transport (GO:0006814)	33.29	2.65	3.65	0.008	0.140	0.041	3.42	0.048

Table 2. Results summary for the selected candidate genes. [a]In the RNAseq analysis, expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads) and mean difference between groups as the $\log_2$ transformed fold change (Log_2FC). [b]In the RT-qPCR analysis, expression values are shown as mean relative quantities (RQ) and mean difference between groups is represented as the fold change (FC).

responsible for this. Sequencing analysis revealed a total of 341 miRNAs being expressed in H and 329 in L prolificacy samples. Among all expressed microRNAs found in our endometrial samples, a total of 10 mature miRNAs were predicted as differentially expressed between H and L prolificacy phenotypes when considering a $p\text{-}value < 0.05$. However, we lost this significance when applying the same FDR correction significance criteria as used for DEG identification (Supplementary table S4).

To explore the possible regulatory role of these differentially expressed miRNAs, we predicted their potential target genes using TargetScan software. Five of these 10 differentially expressed miRNAs had as a putative mRNA target one of the DEGs found between the H and L groups (Supplementary table S5). The novel prediction tool from the mirDeep package allowed us to also identify 15 putative novel miRNAs in H samples and 12 in L samples, with an estimated probability of being a genuine miRNA precursor greater than 90% (Supplementary table S6).

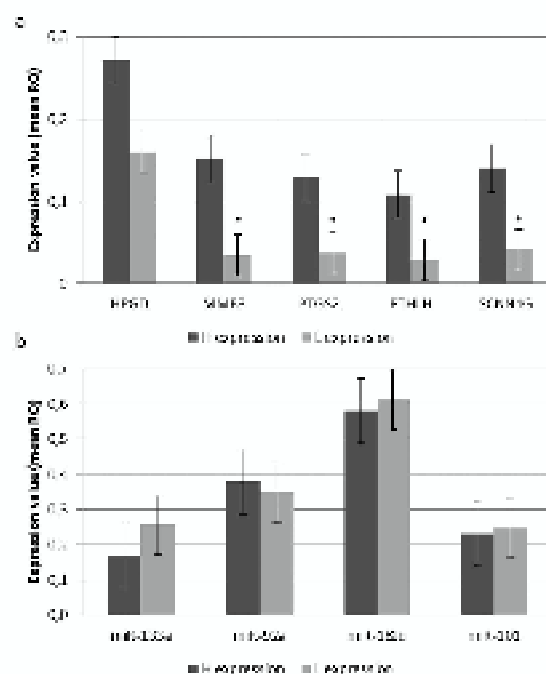


Figure 1. **(a)** RT-qPCR analysis results for gene expression. Expression values were calculated applying the $-2^{\Delta\Delta CT}$ algorithm. Estimated relative quantities were normalized for the expression value of two uterus endogenous genes *B2MG* and *UBC* and calibrated to the sample with a higher expression. Significance was set at a *p-value* < 0.05 (*). **(b)** RT-qPCR analysis results for miRNA expression. Relative quantities were calculated using target-specific amplification efficiencies and normalized for the expression level of two uterus reference miRNAs: *has-miR-93* (M = 0.464; CV = 0.156 and *ssc-miR-103* (M = 0.464; CV = 0.166).

Candidate miRNAs selection and expression levels validation: RT-qPCR. Among the 10 miRNAs found differentially expressed in the RNAseq analysis (*q-value* < 0.05), we selected as candidates those that have been extensively reported in the literature as relevant in the regulation of reproduction-related genes in both pig and human: *ssc-miR-92a*, *ssc-miR-101*, *ssc-miR-133a* and *ssc-miR-181d* (See Table 3). We validated their expression levels by RT-qPCR in the same 36 F_2 extreme individuals (H, n = 18; L, n = 18) used for gene expression validations (Table 3).

Obtained results revealed similar expression levels between both prolificacy groups for these four miRNAs (Fig. 1b). However, significant correlations were found between the expression level of prolificacy-related miRNAs *ssc-miR-92a* and *ssc-miR-133a* and validated DEG analysed by RT-qPCR (Table 4). Again, the observed fold changes were similar in both analysis: *ssc-miR-92a* (RNAseq FC = 1.26, RT-qPCR FC = 1.09), *ssc-miR-101* (RNAseq FC = 1.20, RT-qPCR FC = 0.94), *ssc-miR-181d-5p* (RNAseq FC = 1.16, RT-qPCR FC = 0.95). This confers consistency to our findings and led us to think that the observed differences in the expression levels between H and L groups represent the real biological background of our samples.

Biological role of candidate genes: Interactions and upstream regulators. To place the results in a biological context that allows us to better understand them, we performed an Ingenuity Pathway Analysis (IPA) to analyze the existing networks and potential molecular interactions between the validated candidate genes. Along pregnancy, hormones and other molecules secreted from the porcine conceptus act directly on the endometrium promoting its interaction with maternal uterus and placental development. We identified multiple links and interactions between our validated candidate genes and some molecular components. In the predicted network generated by IPA algorithm (Fig. 2), we observed that the expression of our four validated candidates could be modulated mainly by three molecules: trypsin (for genes *MMP8*, *PTGS2* and *SCNN1G*), insulin (for gene *SCNN1G*) and the vascular endothelial growth factor (*Vegf*) which acts on *PTHLH* gene.

After performing the analysis of the putative common upstream regulators we identified that the common regulators to all four genes are the cytokines Interleukin 1 beta (*ILK-1β*, *p-value* = 0.000007) and the tumor necrosis factor ligand (*TNF*, *p-value* = 0.00008). Results are shown in Table 5.

Discussion. In this study, we investigated the whole transcriptome profile of the swine endometrial epithelium in an Iberian x Meishan F_2 population using RNA sequencing (RNA-seq), with the aim to

miRNA	Function	References	DEG predicted as target	RNAseq analysis[a] High (RPKM)	RNAseq analysis[a] Low (RPKM)	Log2 FC	*q-value*	RT-qPCR analysis[b] High (RQ)	RT-qPCR analysis[b] Low (RQ)	FC	*q-value*
ssc-miR-92a	Angiogenesis	Bellera N. *et al.*, 2014	*HPGD*	51,874.13	21,710.41	−1.26	0.032	0.376	0.347	1.09	0.515
	Embryo implantation	Su L. *et al.*, 2014									
	Placentation	Su L. *et al.*, 2014									
	trophoblast differentiation	Kumar P. *et al.*, 2013									
ssc-miR-101	Ginecological tumors	Torres A. *et al.*, 2010	*HTRA3, ATP1B1, PTGS2, JUNB*	430.21	187,19	−1.20	0.034	0.231	0.247	0.94	0.829
	Embryo implantation	Chakrabarty A. *et al.*, 2007									
	Endometriosis	Teague E. *et al.*, 2010									
ssc-miR-133a	Uterine tumors	Torres A. *et al.*, 2010	*ENPEP*	533.76	1,777.17	1.74	0.050	0.168	0.255	0.66	0.290
	Skeletal muscle development	Lee J. *et al.*, 2013									
ssc-miR-181d	Hypoxia	Shen G. *et al.*, 2013	*MMP8, MME*	55.51	124.35	1.16	0.046	0.580	0.611	0.95	0.698
	Embryo implantation	Su L. *et al.*, 2014									
	Placentation	Su L. *et al.*, 2014									
	Endometrial stromal decidualization	Estella C. *et al.*, 2012									

Table 3. Results summary for the validated candidate miRNAs. [a]In the RNAseq analysis, expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads) and mean difference between groups as the $\log_2$ transformed fold change (Log_2FC). [b]In the RT-qPCR analysis, expression values are shown as mean relative quantities (RQ) and mean difference between groups is represented as the fold change (FC).

		MMP8	*PTGS2*	*PTHLH*	*SCNN1G*
ssc-miR-133a	Pearson's correl.	−0.575	−0.537	−0.533	−0.516
	p-value	0.0003*	0.0007*	0.0008*	0.0013*
	N	36	36	36	36
ssc-miR-181d	Pearson's correl.	−0.140	−0.139	−0.088	−0.137
	p-value	0.4159	0.4199	0.6113	0.4240
	N	36	36	36	36
ssc-miR-101	Pearson's correl.	−0.059	−0.045	−0.069	−0.123
	p-value	0.7380	0.7889	0.6925	0.4824
	N	35	35	35	35
ssc-miR-92a	Pearson's correl.	0.630	0.574	0.615	0.551
	p-value	0.00004*	0.0002*	0.00007*	0.0005*
	N	36	36	36	36

Table 4. Pearson's correlations between miRNA expression values obtained by RT-qPCR and validated target genes expression. Significance was set at a *p-value* < 0.05. (*)Correlation is significant at the 0.01 level (bilateral).

identify key differences in gene expression associated to swine reproductive efficiency. Understanding the complexity of the key mechanisms for successful reproduction in humans and animals has been challenging. Even though a few studies have addressed this goal, this study represents one of the first

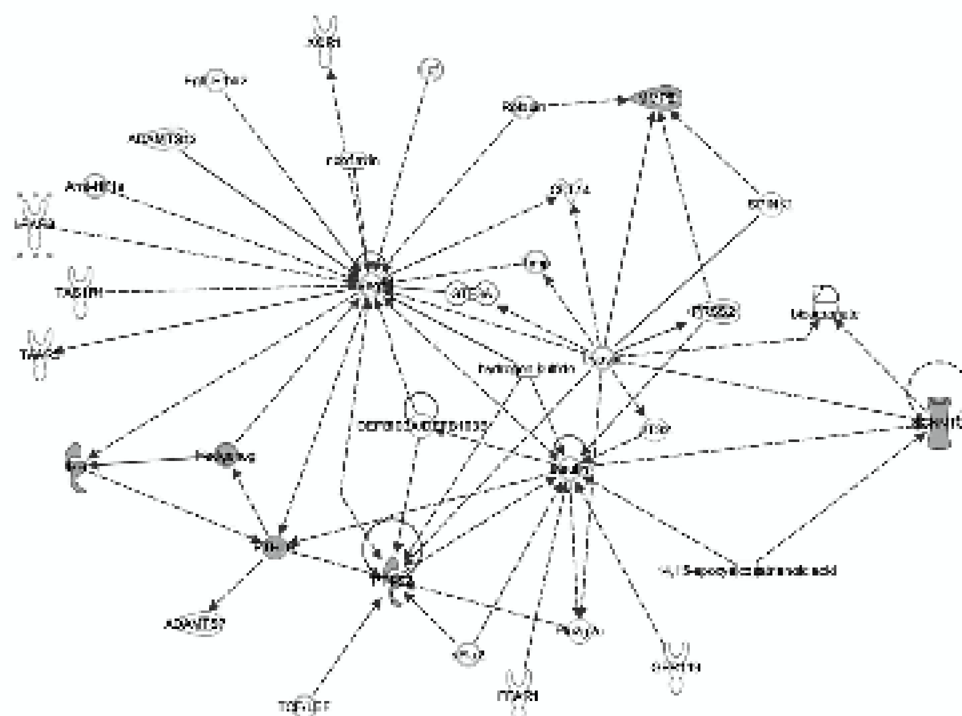


Figure 2. Ingenuity Pathway Analysis (IPA) Core Analysis-based network. Links of validated genes and other genes or molecules are represented with a continuous (direct interaction) or discontinuous line (indirect interaction).

Upstream Regulator	Molecule Type	*p*-value	Target molecules
Dexamethasone	Chemical drug	0.000001	*MMP8,PTGS2,PTHLH,SCNN1G*
IL1B	Cytokine	0.000007	*MMP8,PTGS2,PTHLH,SCNN1G*
TNF	Cytokine	0.000079	*MMP8,PTGS2,PTHLH,SCNN1G*
Lipopolysaccharide	Chemical drug	0.000097	*MMP8,PTGS2,PTHLH,SCNN1G*

Table 5. Network associations of upstream regulators and validated candidate genes predicted by Ingenuity Pathway Analysis (IPA). The Core Analysis calculates the predicted upstream regulators based on the FC direction (up-regulated or down-regulated) observed among known downstream targets.

descriptions of the mechanisms that affect embryonic survival in the pig, providing the knowledge to enhance fertility and reproductive health in this species.

The main limitation of increasing litter size in pigs is prenatal mortality. Two critical stages are early and mid-gestation, responsible for around 20–30% (days 10–30 of gestation) and 10–15% (days 50–70 of gestation) of embryonic loss respectively[2]. Recent evidences have indicated that the prenatal loss in pigs results mainly from the decreased placental efficiency and uterine capacity[33,34].

Uterine receptivity to implantation is a process that can be very different, depending on the species, but always involves several changes in the expression of genes that are directly involved in pathways, such as progesterone and oestrogen biosynthesis, immune recognition, membrane permeability, angiogenesis and vasculogenesis, transport of nutrients and signalling for pregnancy recognition[35,36]. Thus, changes in the expression level of those genes may influence uterine receptivity to implantation.

In this study we have identified 141 differentially expressed genes between high and low prolificacy samples. Functional enrichment analysis suggested that most of these genes are directly involved in the above-mentioned biological processes, which are highly relevant for pregnancy and some specific stages of embryonic development in swine. We have focused our validations on a first set of genes that are up-regulated in our high-prolificacy samples. Some of those genes are also located inside the confidence intervals of previously described reproduction QTLs: ovulation rate, gestation length, number of piglets born alive and embryo's birth weight. Considering these, we proceeded to validate their expression by real time RT-qPCR. As predicted in the RNAseq analysis, four of these genes were differentially expressed in our endometrial samples, being overexpressed in those with a high-prolificacy phenotype.

Several DEGs found in our samples have been extensively discussed by many authors before[37–42], and their involvement in the establishment of pregnancy and in the physiological, molecular and structural changes that take place in the uterine tissue to promote embryo implantation have been demonstrated in pigs and other mammals . Their involvement in many stages of embryonic development postulate them as key factors for deciphering the mechanisms involved in the regulation of litter size in our study population.

Prostaglandins (PGs) produced by the uterus play an important role in regulation of the oestrous cycle and during early pregnancy in pigs and many other species[43]. In the porcine endometrium, luteoprotective *PGE2* and luteolytic *PGF2α* are the main PGs produced and pregnancy establishment depends directly in a proper ratio between the synthesis of both. An inhibition of PG synthesis results in pregnancy failure[44]. One of the validated genes found differentially expressed among our samples is the prostaglandin endoperoxide synthase (*PTGS*; also known as prostaglandin G/H synthase or cyclooxygenase *COX2*). The *PTGS2* gene has been widely discussed over the years and its key function to ensure reproductive success has been widely demonstrated through several previous studies. It constitutes a rate-limiting enzyme in the production of PGs as it catalyzes the conversion of arachidonic acid to *PGH2*, which is a common substrate for various prostaglandins. Its conserved role in implantation in various species, including humans, has previously been discussed[45,46]. Thus, considering that the production of prostaglandins directly contributes to the successful establishment of pregnancy, and that uterine receptivity to implantation is progesterone-dependent, a lack in the expression of this gene will directly affect the appropriate conceptus attachment. It has been observed that the expression of *PTGS1* and *PTGS2* is substantially increased during implantation. We speculate that the underexpression of this gene in our low-prolificacy samples may contribute to embryonic deaths due to deficiencies in progesterone synthesis. This uterine receptivity via expression of *PTGS2* gene is a process that has been demonstrated to be directly regulated by another key gene also found DE in our samples: *KLF5*. This gene belongs to the Kruppel-like factors (KLFs) family. This is a zinc finger-containing transcription factor, which is known to regulate several cellular processes, including development, differentiation, proliferation, and apoptosis[47]. At the beginning of the attachment reaction, the first cell type to interact with the blastocyst trophectoderm is the uterine luminal epithelium. *KLF5* function is critical to make this uterine luminal epithelium conducive to blastocyst implantation and growth. In its absence, trophectoderm development is defective, resulting in developmental arrest at the blastocyst stage[48]. These results suggest that *KLF5* is a key regulator of embryo pre-implantation[49]. Thus, the fact that this gene is overexpressed in our high-prolificacy samples strengthens our idea of the important effect it may have on prolificacy levels and litter size control.

As mentioned before, successful establishment of pregnancy also depends on many structural changes that take place in the uterine tissue. Species with invasive implantation require a cell-to-cell communication through connexin proteins. Although porcine implantation is superficial, some authors have reported that endometrial cell-to-cell interaction may also be necessary for limiting trophoblast invasiveness or to develop specific channels that allow this superficial implantation[50]. And it is at this stage where the validated gene *MMP8* plays a key role. Proteins such as matrix metalloproteinase (MMP) are a family of enzymes (with more than 20 members identified) that use zinc-dependent catalysis to break down the components of the extracellular matrix (ECM)[39,51,52]. We hypothesize that the observed significant overexpression of this gene in our high-prolificacy samples may indicate a more efficient tissue reorganization to support the growing foetus.

Another relevant structural gene found differentially expressed in our extreme F_2 population is the Forkhead transcription factor *FOXA2*. *FOXA* transcription factors are a subfamily of Forkhead transcription factors that has been found to play an important role in early development, organogenesis, metabolism and homeostasis[53]. Low-prolificacy samples show a decreased expression of this gene compared to those with high prolificacy, supporting our idea that an underexpression of this gene could be leading to defects in early development, affecting stages such as gastrulation or, later on, in embryo morphogenesis.

Many other genes found differentially expressed in this study are closely related with critical stages in embryo development at implantation level or later in the survival of the embryo itself. This has provided us with a powerful list of candidates that require further validations in order to prove their direct involvement in the control of litter size in swine. Because of the usefulness of the pig as a biomedical model and the parallelism in the function of these genes in humans, this study also provides a powerful tool to understand which genes are key in the process of embryo survival in mammals.

We also wanted to explore the regulatory mechanisms that do mediate this differential expression in our study population. To do so, we have also analysed the miRNA expression profile in both extreme phenotypic groups.

We predicted a differential expression of 10 mature miRNAs between our H and L prolificacy samples. Some of these differentially expressed miRNAs have been demonstrated to be directly involved in the regulation of reproductive-related genes in pig and other mammals[54–57]. After this preliminary bioinformatic screening we proceeded to the experimental validation of the expression level of 4 of these 10 miRNAs, considering their role in reproductive-related pathways: *ssc-miR-92a*, *ssc-miR-101*, *ssc-miR-133a* and *ssc-miR-181d*.

In concordance with RNAseq predictions, *ssc-miR-101*, *ssc-miR-133a* and *ssc-miR-181d* were overexpressed in L samples while *ssc-miR-92a* was overexpressed in H samples. *MiR-92*, belongs to the

miR-17~92 cluster, demonstrated in recent reports to regulate cardiac development, endothelial cell proliferation and angiogenesis, which are relevant processes for embryogenesis and pregnancy itself[58]. Loss and gain of function experiments showed that miR-92a inhibited angiogenesis *in vitro* and *in vivo*[59] and that deletion of *miR-92a* is sufficient to induce a developmental skeletal defect[55]. Thus, the observed overexpression of this miRNA in our H samples could be explained by its positive effect on several key processes for pregnancy and embryo development.

Real-time RT-qPCR analysis revealed similar expression levels of these miRNAs in both groups (FC < 1.5). However, it has been demonstrated that even very small changes in microRNA expression levels (FC 1.5 to 2.5) could have a direct impact on their target genes and some authors have observed these small differences when performing miRNA differential expression studies related to reproductive processes[60,61]. We hypothesize that this could be caused by an insufficient sequencing depth in our libraries, because despite these similar miRNA expression levels observed between both phenotypes, a significant correlation was found between the expression levels of validated genes *PTHLH*, *MMP8*, *PTGS2* and *SCNN1G*, and both *ssc-miR-133a* and *ssc-miR-92a*. Therefore, the finding of this significant correlation leads us to think that the observed differences, despite being low, may be biologically significant. Many years ago, Calin *et al.* suggested that the capability of miRNAs to regulate multiple targets within the same pathway could amplify their biological effects[62].

Besides miRNAs, upstream regulators such as transcription factors (TFs), growth factors (GFs) and many other molecules may play a critical role as *drivers* or master regulators of gene expression. Investigating their involvement in a particular gene network or pathway can provide better clues on the underlying regulatory mechanisms that do mediate the observed differences in the expression of key genes in a particular biological context.

In this study we have explored the regulatory role that some candidate miRNAs exert in the expression of key reproductive-related genes and the possible effect that this has on litter size control. In addition, we have established which interactions exist between our validated candidate genes and other known regulatory molecules. There are two cytokines particularly capable of acting on the expression of these four genes which are the *ILK-1β* and the *TNF*.

In reproductive biology, the role of these cytokines has been implicated in ovulation, menstruation, and embryo implantation, and pathological processes such as preterm delivery, and endometriosis[63,64]. The interleukin 1 is a pro-inflammatory cytokine with multiple functions in a range of tissues[65]. All components of the IL-1 system have been examined in the human endometrium and have been implicated as an important mediator of embryo implantation[66,67]. Simón C. and collaborators, demonstrated in mice, that IL-1 receptor antagonist given before implantation significantly reduces the number of implanted embryos, indicating a role for IL-1 in embryo implantation[64].

The TNF is a pro-inflammatory cytokine that plays an important role in modulating the acute phase reaction. It was first discovered in amnion and placenta[68], but many studies have demonstrated the presence of this cytokine and its receptors in the diverse human reproductive tissues[69]. The TNF has been implicated in ovulation, corpus luteum formation and luteolysis, and it has been related to many endometrial and gestational diseases such as amniotic infections, recurrent spontaneous abortions, preeclampsia, preterm labour or endometriosis[70–72]. Although these cytokines may be acting on the expression of our validated candidate genes, we haven't seen them differentially expressed between H and L groups.

It is clear, that there is a complex network of interacting genes regulating litter size in pigs. However, this work has led to the identification of several potential candidate genes associated with critical steps involved in embryonic survival during the sow's gestation. Our results also describe the possible regulatory mechanisms that could be responsible of the differences in the expression level of key genes related with litter size control in pigs.

Materials and Methods

Animal material and sample collection. Animals used in this study come from an F_2 population resulting from the cross of 3 Iberian males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) with 18 Meishan females (Domaine du Magneraud, INRA, France). Once the F_1 generation was obtained, 8 boars and 97 sows were mated to obtain a 255 F_2 progeny at the Nova Genètica S.A. experimental farm (Lleida, Spain).

During four consecutive parities, main parameters based on the sows' reproductive efficiency were recorded: number of piglets born alive (NBA) and total number of piglets born (TNB) means. At day 30–32 of their fifth gestation, when litter size has reached the maximum[73], sows were slaughtered and the number of *corpora lutea* (CL or OR) and number of foetuses (NF) attached to the uterus were also recorded. At slaughter, endometrial samples from the apical uterus of F_2 sows were collected and subsequently snap-frozen in liquid nitrogen. Preservation and storage was made at −80 °C until usage. All animal procedures were carried out according to the European animal experimentation ethics law and approved by the institutional animal ethics committee of Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Phenotypic records and samples selection. F_2 sows were ranked by their estimated breeding value (EBV), which was calculated by using best linear unbiased predictors (BLUP) according to the

reproductive traits described above: NBA and TNB means, OR and NF. Based on this ranking, individuals were divided into two groups: high (H; EBV > 0) and low (L; EBV < 0) prolificacy. Among the whole F_2 progeny (n = 255), individuals displaying the most extreme EBVs were selected to be used in this study (n = 36). All phenotypic records are shown in Table 6.

RNA isolation and quality assessment. Total RNA was extracted from sows' endometrial samples using TRIzol® reagent (Invitrogen, Carlsbad, USA), following the manufacturer's instructions. The RNA integrity was assessed using an Eukaryote Total RNA Nano 6000 Labchip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA) and quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA). Only those RNA samples with an RNA integrity number (RIN) ≥ 7 were used in subsequent experiments.

Ion Torrent PGM libraries preparation and RNA sequencing. Ion Torrent adapter-ligated libraries were prepared from extracted total RNA according to the Ion Total RNA-seq Kit v2 protocol (Life Technologies – Part #4476286 Rev. B) following the manufacturer's instructions.

mRNA libraries preparation. Samples corresponding to animals displaying very extreme EBVs and very high RNA quality (RIN≥ 8) were used to prepare mRNA libraries (H, n = 3; L, n = 3). We constructed sequencing libraries starting from 500 ng of total RNA. PolyA RNA fraction was purified from total RNA samples using the Dynabeads® mRNA DIRECT Micro Kit (Life Technologies – Part #1148804 Rev. A) following the manufacturer's instructions. Each sample was subjected to Ion semiconductor sequencing using a 318 chip on an Ion-Torrent PGM sequencer.

Small RNA libraries preparation. Small RNA sequencing was also performed using 318 chips on an Ion-Torrent PGM sequencer. In this case, we used stored GS FLX 454 microRNA sequencing libraries that we had previously used in our research[61], which included the same extreme samples used in the mRNA libraries protocol (H, n = 7; L, n = 5). To adapt these performed libraries to the Ion semiconductor sequencing technology protocol, it was necessary to remove the 454 specific adaptors and to add the Ion Torrent A and P1 specific ligators. After doing so, each miRNA library was re-sequenced.

Bioinformatics and statistical analysis. Approximately 5 million short single-end reads (≈200 bp) were obtained for each library and sample and were subsequently assembled into a non-redundant set of 30,585 gene transcripts (3,024,654,544 bp) from the available *Sus scrofa* genome alignment version 10.2 (available at http://www.ncbi.nlm.nih.gov/assembly/GCF_000003025.5/#/def). In average, 75% of the reads were successfully mapped to the *Sus scrofa* genome.

Quality control for single-end raw reads. Raw reads formatted as fastq files were processed using FastQC 0.10.1 (freely available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Considered low quality reads by applying FastQC defaults, were removed and all downstream analyses were performed only on those reads meeting the quality criteria. Ion Torrent A and P1 adaptors were removed using Cutadap 1.4 (freely available at http://code.google.com/p/cutadapt/).

Reads mapping, alignment and annotation. Obtained sequence reads from mRNA libraries were mapped with Tophat (v1.4.0) to the latest porcine genome sequence assembly (Sscrofa10.2, August 2011). Transcript isoforms were assembled using Cufflinks 2.1.1 and combined with gene annotations extracted from Ensembl (ftp://ftp.ensembl.org/pub/release-75/gtf/sus_scrofa). The criteria used to filter out unique sequence reads was: minimum length fraction of 0.9; minimum similarity fraction of 0.8 and a maximum number of 2 mismatches.

Sequence reads from small RNA libraries were analysed following the Perl scripts contained in the miRDeep 2.0 package[74] (freely available at http://www.mdc-berlin.de/rajewsky/miRDeep). Briefly, reads were first collapsed to ensure that each sequence only occurs once. Collapsed reads were then mapped to predefined miRNA precursor sequences from the miRBase v.20 contained in the porcine genome sequence assembly (Sscrofa 10.2, August 2011). Finally, unmapped reads served as input sequences for the novel miRNAs prediction algorithm.

Differential gene expression, functional annotation and QTL mapping analysis. Analysis of differential gene expression across high and low-prolificacy groups was performed using Cuffdiff 2.0.2 which is included in the Cufflinks package (available at http://cufflinks.cbcb.umd.edu/manual.html). For small RNA libraries, differentially expressed miRNA genes were detected by using the DEseq R package 1.8.3[75]. A Benjamini-Hochberg FDR corrected *p-value* of 0.05 was set as the threshold for significant differential expression in both cases.

Babelomics 4.3.0 (http://babelomics.bioinfo.cipf.es) was used to functionally annotate DEG. The pig functional annotation database is not as complete as human, therefore, we converted the pig gene IDs (Ensembl *Sus scrofa* 10.2) into human gene IDs using Ensembl BioMart tool (http://www.ensembl.org/biomart/martview/). Then the homologous human Ensembl IDs were submitted to the Babelomics

Prolificacy level	Animal	NBA[a]	TNB[a]	OR[b]	NF[b]	EBV
HIGH	A1 (791)	12.00	10.00	13.00	10	1.73
	A2 (787)[c,d]	11.75	13.00	16.00	16	1.68
	A3 (169)	12.25	11.00	14.00	11	1.68
	A4 (332)[c,d]	12.75	13.33	16.00	14	1.55
	A5 (373)[c,d]	11.25	11.00	20.00	17	1.50
	A6 (878)[d]	12.00	10.50	14.00	7	1.42
	A7 (425)	11.00	11.00	0.00	13	1.34
	A8 (767)	9.40	10.50	17.00	14	1.31
	A9 (20)	11.00	10.00	20.00	14	1.22
	A10 (127)	11.00	11.67	17.00	13	1.21
	A11 (365)	10.50	10.00	16.00	9	1.17
	A12 (389)[d]	10.25	10.50	19.00	16	1.09
	A13 (597)	10.00	9.50	20.00	11	0.92
	A14 (151)	10.75	12.00	20.00	13	0.89
	A15 (874)[d]	10.25	10.00	11.00	8	0.82
	A16 (271)	10.50	9.67	15.00	14	0.81
	A17 (30)	10.75	10.67	19.00	13	0.80
	A18 (485)	11.00	12.50	16.00	16	0.77
Average (HIGH)		11.02	10.94	15.72	12.72	1.22
LOW	A19 (350)[c,d]	4.50	3.00	15.00	6	−2.48
	A20 (309)	5.00	4.33	16.00	8	−2.42
	A21 (360)[c,d]	5.00	5.33	18.00	1	−2.33
	A22 (260)	4.75	5.00	17.00	10	−2.31
	A23 (173)	5.00	6.67	15.00	10	−2.30
	A24 (861)[c,d]	6.50	6.00	24.00	9	−2.04
	A25 (409)	4.75	5.67	18.00	11	−1.94
	A26 (918)	7.00	8.50	16.00	13	−1.46
	A27 (779)	6.25	5.50	23.00	10	−1.45
	A28 (915)	4.75	4.00	18.00	8	−1.21
	A29 (443)	5.25	6.50	16.00	5	−1.13
	A30 (702)	6.00	7.50	13.00	11	−1.06
	A31 (322)[d]	4.75	5.00	16.00	14	−0.95
	A32 (204)	5.00	3.67	14.00	15	−0.95
	A33 (486)[d]	5.25	3.50	24.00	5	−0.91
	A34 (499)	6.75	6.50	13.00	11	−0.59
	A35 (895)	7.25	8.50	13.00	10	−0.46
	A36 (846)	6.75	5.00	22.00	14	−0.45
Average (LOW)		5.53	5.51	17.28	9.50	−1.47

Table 6. Phenotypic records of the F_2 Iberian × Meishan sows used in this study. [a]NBA (number of piglets born alive) and TNB (total number of piglets born) trait entries correspond to the average for four consecutive parities. [b]OR (number of *corpora lutea*) and NF (number of foetuses) recorded at slaughter on the fifth gestation. [c]Extreme samples used for mRNA libraries preparation and sequencing. [d]Extreme samples used for microRNA libraries sequencing.

database for functional annotation. *P-values* to estimate over-represented GO terms were obtained through Chi square analysis. An FDR-corrected *p-value* of 0.10 was set as the threshold for significance.

All differentially expressed genes found were mapped against the latest release (Aug 25, 2014) of the Pig Quantitative Trait Locus Database[76]. Those DEGs displaying a significant functional annotation related to reproduction processes and/or a positive mapping into known reproductive QTLs were selected as a first set of candidates for quantitative real-time PCR validations.

Gene	Forward Primer	Reverse Primer	Type	Conc.
B2MG	ACCTTCTGGTCCACACTGAGTTC	GGTCTCGATCCCACTTAACTATCTTG	Endogenous	300nM
HPGD	CAGGCACAACTTAGAGATACATTTAGG	TCCAGCATTATTGACCAAAATGTC	Target gene	300nM
MMP8	GGACCAAAAACCTCCAAAAATTACA	TGAGACAGCCCCAAGGAATG	Target gene	300nM
PTGS2	ACGAGCAGGCTGATACTGATAGG	GTGGTAGCCACTCAGGTGTTGTAC	Target gene	300nM
PTHLH	GCCGCCGACTCAAAAGAG	CGCCGTAAATCTTGGATGGA	Target gene	300nM
SCNN1G	GCTGCCTACTCCCTGCAGATC	TACTGAGCGCACCCACATTTC	Target gene	300nM
UBC	GCATTGTTGGCGGTTTCG	AGACGCTGTGAAGCCAATCA	Endogenous	300nM

Table 7. Primers used for the genes RT-qPCR validation design.

Expression level validation by reverse transcription quantitative real-time PCR (RT-qPCR). Five candidate genes and four candidate miRNAs displaying significant differences in their expression level between H and L samples were validated by RT-qPCR. The same samples selected for RNA-seq were used in these validations, but in order to obtain a broader view of the expression level of these genes in our population, the sample size was expanded using other extreme F_2 samples (H, n = 18; L, n = 18).

Reverse transcription (RT): cDNA synthesis. Extracted total RNA was quantified using an ND 1000 Nanodrop® Spectrophotometer (Thermo Scientific, Wilmington, USA). The RNA quality and integrity were determined using an Eukaryote Total RNA Nano 6000 Labchip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA).

Synthesis of cDNA for gene expression validation was performed using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) from 1 μg of total RNA in 20 μl reaction. The synthesis of cDNA for miRNA expression validation was performed using extracted total RNA as described by Balcells *et al.*[77] Briefly, 600 ng of total RNA in a final volume of 20 μL including 10x poly (A) polymerase buffer, 0.1 mM of ATP, 0.1 mM of each dNTP, 1 μM of RT-primer, 200 U of M-MuLV Reverse Transcriptase (New England Biolabs, USA) and 2 U of poly (A) polymerase (New England Biolabs, USA) was incubated at 42 °C for 1 hour and 95 °C for 5 minutes for enzyme inactivation. The used RT-primer sequence was 5′-CAGGTCCAGTTTTTTTTTTTTTTTVN, where V is A, C and G and N is A, C, G, and T. Minus reverse transcription (RT) and minus poly A) polymerase controls were performed.

Real-time RT-qPCR reaction. DE genes expression validation. Quantitative PCR reactions were performed in triplicate in 20 μL final volume including 10 μL SYBR® Select Master Mix (Life Technologies – Thermo Fisher Scientific, Massachusetts, USA), 300 nM of each primer and 5 μL of a 1:200 dilution of the cDNA. A 1:5 relative standard curve generated from a pool of equal amounts of cDNA from all samples was included in each qPCR assay to estimate qPCR efficiency. Reactions were incubated in a 96-well plate at 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 sec and 60 °C for 1 min on a 7900 HT Real-Time PCR System using 7900HT SDS v2.4 software (Applied Biosystems, USA). DNA primers for each gene were designed using Primer Express® software v2.0 (Applied Biosystems, USA) following manufacturer's instructions (Table 7). Melting curve analysis was included in each qPCR to detect unspecific amplifications. Expression values were calculated with qbasePLUS software (Biogazelle) applying the $-2^{\Delta\Delta Ct}$ algorithm, after verifying that the assumptions of the method were met[78]. Estimated relative quantities were calibrated to the sample with a higher expression and normalized for the expression value of two uterus endogenous genes: *B2MG*[79] and *UBC*[80]. Reference genes stability was also assessed with qBasePLUS software considering a GeNorm M value < 0.5 and a coefficient of variation (CV) < 0.2. Significance was set at a *p-value* < 0.05.

DE miRNAs expression validation and putative targets prediction. Quantitative PCR reactions were performed as described above but using a different concentration of primers according to each miRNA. DNA primers were designed following the methodology suggested by Balcells *et al.* (Table 8). Relative standard curves were included in each qPCR assay to estimate target-specific amplification efficiencies. Expression values were calculated with qbasePLUS software using these amplification efficiencies. Relative quantities were normalized for the expression value of two uterus reference miRNAs: *has-miR-93* and *ssc-miR-103*[81] and calibrated to the sample with a higher expression. Reference miRNAs stability was determined considering a GeNorm M value < 0.5 and a coefficient of variation (CV) < 0.2. Significance was set at a *p-value* < 0.05.

Biological putative targets prediction was performed using TargetScan 6.2 online software. Targets were considered true positives if conserved 8mer and 7mer sites match the seed region of each miRNA.

Analysis of candidate genes interactions and upstream regulators. The four validated genes (*MMP8*, *PTHLH*, *PTGS2* and *SCNN1G*) were submitted to Ingenuity Pathway Analysis (IPA 4.0, Ingenuity Systems Inc., www.ingenuity.com) for mapping to canonical pathways and identifying

miRNA	miRNA Sequence	Forward Primer	Reverse Primer	Conc.
hsa-miR-93	CAAAGTGCTGTTCGTGCAGGTAG	GCAAAGTGCTGTTCGTG	TCCAGTTTTTTTTTTTTTTTTTCTACCT	200 nM
ssc-miR-92a-2	TATTGCACTTGTCCCGGCCTGT	AGGTGTGTATAAAGTATTGCACTTGTCC	CAGGTCCAGTTTTTTTTTTTTTTTTACAG	250 nM
ssc-miR-101-1	TACAGTACTGTGATAACTGAA	GCTGTATATCTGAAAGGTACAGTACTGTGAT	GGTCCAGTTTTTTTTTTTTTTTTTCAGTT	250 nM
ssc-miR-103	AGCAGCATTGTACAGGGCTATGA	AGAGCAGCATTGTACAGG	GGTCCAGTTTTTTTTTTTTTTTTTCATAG	250 nM
ssc-miR-133a-1	TTGGTCCCCTTCAACCAGCTG	GAATGGATTTGGTCCCCTTCA	CAGTTTTTTTTTTTTTTTTTCAGCTGGT	250 nM
ssc-miR-181d-5p	AACATTCAACGCTGTCGGTGAGTT	CACAATCAACATTCATTGTTGTCG	TCCAGTTTTTTTTTTTTTTTTTTTAACCCAC	250 nM

Table 8. Primers used for the miRNAs RT-qPCR validation design.

upstream regulators. As the Ingenuity Knowledge Base relies on ortholog information for only human, mouse, and rat, we submitted to IPA the correspondent human Ensembl IDs of our candidate genes. We ran the Core Analysis function designating a set of criteria: genes and endogenous chemicals, direct and indirect interactions, maximum molecules per network (35) and networks per analysis (25), humans as the selected specie, all tissues and primary cells. The resulting networks were scored based on the fold change provided by Cuffdiff as log_2 (fold change) for each gene. The obtained *p-values* correspond to the Fisher's exact test, with the null hypothesis that the molecules within the networks are connected based on chance.

References

1. Rydhmer, L. Genetics of sow reproduction, including puberty, oestrus, pregnancy, farrowing and lactation. *Livest. Prod. Sci.* **66**, 1–12 (2000).
2. Spötter, A. & Distl, O. Genetic approaches to the improvement of fertility traits in the pig. *Vet. J.* **172**, 234–47 (2006).
3. Pope, W. F. & First, N. L. Factors affecting the survival of pig embryos. *Theriogenology* **23**, 91–105 (1985).
4. Gu, T. *et al.* Endometrial gene expression profiling in pregnant Meishan and Yorkshire pigs on day 12 of gestation. *BMC Genomics* **15**, 156 (2014).
5. Haley, C. S. & Lee, G. J. Genetic basis of prolificacy in Meishan pigs. *J. Reprod. Fertil. Suppl.* **48**, 247–59 (1993).
6. Noguera, J. L. *et al.* A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. *BMC Genomics* **10**, 636 (2009).
7. Rothschild, M. F., Hu, Z. & Jiang, Z. Advances in QTL mapping in pigs. *Int. J. Biol. Sci.* **3**, 192–7 (2007).
8. Rathje, T. A., Rohrer, G. A. & Johnson, R. K. Evidence for quantitative trait loci affecting ovulation rate in pigs. *J. Anim. Sci.* **75**, 1486–94 (1997).
9. Rohrer, G. A. *et al.* A comprehensive map of the porcine genome. *Genome Res.* **6**, 371–91 (1996).
10. Wilkie, P. J. *et al.* A genomic scan of porcine reproductive traits reveals possible quantitative trait loci (QTLs) for number of corpora lutea. *Mamm. Genome* **10**, 573–8 (1999).
11. Campbell, E. M. G., Nonneman, D. & Rohrer, G. A. Fine mapping a quantitative trait locus affecting ovulation rate in swine on chromosome 8. *J. Anim. Sci.* **81**, 1706–14 (2003).
12. De Koning, D. J. *et al.* Detection and characterization of quantitative trait loci for meat quality traits in pigs. *J. Anim. Sci.* **79**, 2812–9 (2001).
13. King, A. H., Jiang, Z., Gibson, J. P., Haley, C. S. & Archibald, A. L. Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8. *Biol. Reprod.* **68**, 2172–9 (2003).
14. Cassady, J. P. *et al.* Identification of quantitative trait loci affecting reproduction in pigs. *J. Anim. Sci.* **79**, 623–33 (2001).
15. Rohrer, G. A., Ford, J. J., Wise, T. H., Vallet, J. L. & Christenson, R. K. Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. *J. Anim. Sci.* **77**, 1385–91 (1999).
16. Rothschild, M. F. Genetics and reproduction in the pig. *Anim. Reprod. Sci.* **42**, 143–151 (1996).
17. Samborski, A., Graf, A., Krebs, S., Kessler, B. & Bauersachs, S. Deep sequencing of the porcine endometrial transcriptome on day 14 of pregnancy. *Biol. Reprod.* **88**, 84 (2013).
18. Bauersachs, S. & Wolf, E. Transcriptome analyses of bovine, porcine and equine endometrium during the pre-implantation phase. *Anim. Reprod. Sci.* **134**, 84–94 (2011).
19. Franczak, A., Wojciechowicz, B. & Kotwica, G. Transcriptomic analysis of the porcine endometrium during early pregnancy and the estrous cycle. *Reprod. Biol.* **13**, 229–37 (2013).
20. Liu, X., Xu, Z., Zhu, L., Liao, S. & Guo, W. Transcriptome Analysis of Porcine Thymus following Porcine Cytomegalovirus Infection. *PLoS One* **9**, *e113921* (2014).
21. Siqueira, F. M. *et al.* Unravelling the transcriptome profile of the Swine respiratory tract mycoplasmas. *PLoS One* **9**, e110327 (2014).
22. Ramayo-Caldas, Y. *et al.* Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics* **13**, 547 (2012).
23. Corominas, J. *et al.* Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics* **14**, 843 (2013).
24. Puig-Oliveras, A. *et al.* Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. *PLoS One* **9**, e99720 (2014).
25. Gu, T. *et al.* Endometrial gene expression profiling in pregnant Meishan and Yorkshire pigs on day 12 of gestation. *BMC Genomics* **15**, 156 (2014).
26. Zhang, H. *et al.* Differential gene expression in the endometrium on gestation day 12 provides insight into sow prolificacy. *BMC Genomics* **14**, 45 (2013).
27. Zhou, Q.-Y. *et al.* Detection of differentially expressed genes between Erhualian and Large White placentas on day 75 and 90 of gestation. *BMC Genomics* **10**, 337 (2009).
28. Kotlabova, K., Doucha, J. & Hromadnikova, I. Placental-specific microRNA in maternal circulation–identification of appropriate pregnancy-associated microRNAs with diagnostic potential. *J. Reprod. Immunol.* **89**, 185–91 (2011).

29. Tang, F. *et al.* Maternal microRNAs are essential for mouse zygotic development. *Genes Dev.* **21**, 644–8 (2007).
30. Yu, Z., Jian, Z., Shen, S.-H., Purisima, E. & Wang, E. Global analysis of microRNA target gene expression reveals that miRNA targets are lower expressed in mature mouse and Drosophila tissues than in the embryos. *Nucleic Acids Res.* **35**, 152–64 (2007).
31. Bidanel, J. Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. III. Dominance and epistatic components of heterosis on reproductive traits. *Genet. Sel. Evol.* **25**, 263 (1993).
32. Silió L., R. J. & T. M. A. in *Porc. ibérico Asp. claves* (Buxadé, C. & Daza, A.) 125–149 (S.A. Mundi-Prensa Libros, 2001).
33. Vallet, J. L. & Freking, B. A. Differences in placental structure during gestation associated with large and small pig fetuses. *J. Anim. Sci.* **85**, 3267–75 (2007).
34. Ford, S. P., Vonnahme, K. A. & Wilson, M. E. Uterine capacity in the pig reflects a combination of uterine environment and conceptus genotype effects. *J. Anim. Sci.* **80**, E66–E73 (2001).
35. Bazer, F. W. *et al.* Novel pathways for implantation and establishment and maintenance of pregnancy in mammals. *Mol. Hum. Reprod.* **16**, 135–52 (2010).
36. Cha, J., Sun, X. & Dey, S. K. Mechanisms of implantation: strategies for successful pregnancy. *Nat. Med.* **18**, 1754–67 (2012).
37. Chen, X. *et al.* Differential Gene Expression in Uterine Endometrium During Implantation in Pigs. *Biol. Reprod.* **92**, 52 (2014).
38. Martinez-Giner, M. *et al.* Expression study on the porcine PTHLH gene and its relationship with sow teat number. *J. Anim. Breed. Genet.* **128**, 344–53 (2011).
39. Mousa, A. a *et al.* Preeclampsia is associated with alterations in DNA methylation of genes involved in collagen metabolism. *Am. J. Pathol.* **181**, 1455–63 (2012).
40. Fan, Y.-F., Hou, Z.-C., Yi, G.-Q., Xu, G.-Y. & Yang, N. The sodium channel gene family is specifically expressed in hen uterus and associated with eggshell quality traits. *BMC Genet.* **14**, 90 (2013).
41. Ashworth, M. D. *et al.* Expression of porcine endometrial prostaglandin synthase during the estrous cycle and early pregnancy, and following endocrine disruption of pregnancy. *Biol. Reprod.* **74**, 1007–15 (2006).
42. Blitek, A., Morawska, E., Kiewisz, J. & Ziecik, A. J. Effect of conceptus secretions on HOXA10 and PTGS2 gene expression, and PGE2 release in co-cultured luminal epithelial and stromal cells of the porcine endometrium at the time of early implantation. *Theriogenology* **76**, 954–66 (2011).
43. Blitek, A. *et al.* Expression of cyclooxygenase-1 and -2 in the porcine endometrium during the oestrous cycle and early pregnancy. *Reprod. Domest. Anim.* **41**, 251–7 (2006).
44. Spencer, T. E., Burghardt, R. C., Johnson, G. A. & Bazer, F. W. Conceptus signals for establishment and maintenance of pregnancy. *Anim. Reprod. Sci.* **82-83**, 537–50 (2004).
45. Dey, S. K. *et al.* Molecular cues to implantation. *Endocr. Rev.* **25**, 341–73 (2004).
46. Zhang, S. *et al.* Deciphering the molecular basis of uterine receptivity. *Mol. Reprod. Dev.* **80**, 8–21 (2013).
47. Parisi, S. & Russo, T. Regulatory role of Klf5 in early mouse development and in embryonic stem cells. *Vitam. Horm.* **87**, 381–97 (2011).
48. Sun, X. *et al.* Kruppel-like factor 5 (KLF5) is critical for conferring uterine receptivity to implantation. *Proc. Natl. Acad. Sci. USA* **109**, 1145–50 (2012).
49. Lin, S.-C. J., Wani, M. A., Whitsett, J. A. & Wells, J. M. Klf5 regulates lineage formation in the pre-implantation mouse embryo. *Development* **137**, 3953–63 (2010).
50. Wu, D. *et al.* Molecular characterization and identification of the E2/P4 response element in the porcine HOXA10 gene. *Mol. Cell. Biochem.* **374**, 213–22 (2013).
51. Schäfers, M., Schober, O. & Hermann, S. Matrix-metalloproteinases as imaging targets for inflammatory activity in atherosclerotic plaques. *J. Nucl. Med.* **51**, 663–6 (2010).
52. Wang, H. *et al.* Functionally significant SNP MMP8 promoter haplotypes and preterm premature rupture of membranes (PPROM). *Hum. Mol. Genet.* **13**, 2659–69 (2004).
53. Friedman, J. R. & Kaestner, K. H. The Foxa family of transcription factors in development and metabolism. *Cell. Mol. Life Sci.* **63**, 2317–28 (2006).
54. Chakrabarty, A. *et al.* MicroRNA regulation of cyclooxygenase-2 during embryo implantation. *Proc. Natl. Acad. Sci. USA* **104**, 15144–9 (2007).
55. Penzkofer, D. *et al.* Phenotypic characterization of miR-92a-/- mice reveals an important function of miR-92a in skeletal development. *PLoS One* **9**, e101153 (2014).
56. Hawkins, S. M. *et al.* Dysregulation of uterine signaling pathways in progesterone receptor-Cre knockout of dicer. *Mol. Endocrinol.* **26**, 1552–66 (2012).
57. Kowalewska, M. *et al.* microRNAs in uterine sarcomas and mixed epithelial-mesenchymal uterine tumors: a preliminary report. *Tumour Biol.* **34**, 2153–60 (2013).
58. Bonauer, A. & Dimmeler, S. The microRNA-17-92 cluster: still a miRacle? *Cell Cycle* **8**, 3866–73 (2009).
59. Bonauer, A. *et al.* MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science* **324**, 1710–3 (2009).
60. 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.* **22**, 1175–82 (2010).
61. Balcells Ortega, I. Deciphering the genetic architecture of prolificacy related traits in an experimental Iberian x Meishan F2 intercross. PhD dissertation, Universitat Autònoma de Barcelona, Department de Ciència Animal i dels Aliments, May (2012).
62. Calin, G. A. & Croce, C. M. MicroRNA-cancer connection: the beginning of a new tale. *Cancer Res.* **66**, 7390–4 (2006).
63. Dominguez, F., Pellicer, A. & Simon, C. The chemokine connection: hormonal and embryonic regulation at the human maternal-embryonic interface–a review. *Placenta* **24** Suppl B, S48–55 (2003).
64. Simón, C., Caballero-Campo, P., García-Velasco, J. A. & Pellicer, A. Potential implications of chemokines in reproductive function: an attractive idea. *J. Reprod. Immunol.* **38**, 169–93 (1998).
65. Dunne, A. & O'Neill, L. A. J. The interleukin-1 receptor/Toll-like receptor superfamily: signal transduction during inflammation and host defense. *Sci. STKE* **2003**, re3 (2003).
66. Rossi, M. *et al.* Identification of genes regulated by interleukin-1beta in human endometrial stromal cells. *Reproduction* **130**, 721–9 (2005).
67. Healy, L. L., Cronin, J. G. & Sheldon, I. M. Endometrial cells sense and react to tissue damage during infection of the bovine endometrium via interleukin 1. *Sci. Rep.* **4**, 7060 (2014).
68. Paradowska, E., Blach-Olszewska, Z. & Gejdel, E. Constitutive and induced cytokine production by human placenta and amniotic membrane at term. *Placenta* **18**, 441–6 (1997).
69. Szarka, A., Rigó, J., Lázár, L., Beko, G. & Molvarec, A. Circulating cytokines, chemokines and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex suspension array. *BMC Immunol.* **11**, 59 (2010).
70. Khan, D. A., Ansari, W. M. & Khan, F. A. Pro/Anti-Inflammatory Cytokines in the Pathogenesis of Premature Coronary Artery Disease. *J. Interferon Cytokine Res.* **31**, 561-567 (2011).
71. Haider, S. & Knöfler, M. Human tumour necrosis factor: physiological and pathological roles in placenta and endometrium. *Placenta* **30**, 111–23 (2009).

72. Hecht, J. L. *et al.* Relationship Between Neonatal Blood Protein Concentrations and Placenta Histologic Characteristics in Extremely Low GA Newborns. *Pediatr. Res.* **69**, 68–73 (2011).
73. Hughes, P. E. & Varley, M. A. *Reproduction in the pig*. (Butterworth-Heinemann Ltd, 1980).
74. Friedländer, M. R. *et al.* Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **26**, 407–15 (2008).
75. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
76. Hu, Z.-L., Park, C. A., Wu, X.-L. & Reecy, J M. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res.* **41**, D871–9 (2013).
77. Balcells, I., Cirera, S. & Busk, P. K. Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnol.* **11**, 70 (2011).
78. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–8 (2001).
79. Ferraz, A. L. J. *et al.* Transcriptome architecture across tissues in the pig. *BMC Genomics* **9**, 173 (2008).
80. Martinez-Giner, M., Noguera, J. L., Balcells, I., Fernández-Rodriguez, A. & Pena, R. N. Selection of internal control genes for real-time quantitative PCR in ovary and uterus of sows across pregnancy. *PLoS One* **8**, e66023 (2013).
81. Timoneda, O., Balcells, I., Córdoba, S., Castelló, A. & Sánchez, A. Determination of reference microRNAs for relative quantification in porcine tissues. *PLoS One* **7**, e44413 (2012).

Aknowledgements

This research was funded by the projects AGL2004–08368-C03, AGL2010-22358-C02-01 and by the Consolider-Ingenio 2010 Programme (CSD2007-00036) all from the Spanish Ministry of Economy and Competitiveness. The authors are also indebted to INRA (France) and the CIA El Dehesón del Encinar (Spain) for providing the purebred Meishan sows and Iberian boars, respectively. Funding: This research was funded by the Science and Innovation Ministry of the Spanish Government (projects AGL2010-22358-C02-01 and AGL2004–08368-C03) and also by the Consolider-Ingenio 2010 Programme (CSD2007-00036). SC is a recipient of an FPI PhD fellowship from the Science and Innovation Ministry of the Spanish Government. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Experiments conceived and designed by A.C., A.S. and S.C. Experimental material collection and F_2 management: A.S., C.O. and J.N. Experiments performed by S.C. and A.C. Data analysed by S.C. and A.C. Statistical analysis performed by S.C and A.C. Paper written by S.C. Critical revisions of manuscript: A.C., A.S., C.O., I.B., J.N. and O.T. Funding obtained by AS. Research overview: A.S.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Córdoba, S. *et al.* Endometrial gene expression profile of pregnant sows with extreme phenotypes for reproductive efficiency. *Sci. Rep.* **5**, 14416; doi: 10.1038/srep14416 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

3.1.2. Validation of reproduction-related candidate genes

Key genes for litter size control show significant expression differences in the endometrium of pregnant sows with extreme phenotypes for reproductive efficiency

Córdoba-Terreros, S, Balcells, I., Castelló, A., Ovilo, C., Noguera, J.L., Timoneda, O., Sánchez, A.

Animal Genetics (Subbmitted, August 2015)

Key genes for litter size control show significant expression differences in the endometrium of pregnant sows with extreme phenotypes for reproductive efficiency

Córdoba-Terreros, S.[1*], Balcells, I.[1], Castelló, A.[1], Ovilo, C.[2], Noguera, J.L.[3], Timoneda, O.[1], Sánchez, A.[1]

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

[2] Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (SGIT-INIA), Madrid, 28040, Spain

[3] Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Lleida, 25198, Spain

*Corresponding author:

sarai.cordoba@cragenomica.es

034935636600

ABSTRACT

Background. The annual production of a sow is determined to a large degree by litter size and the capacity of maintaining viable embryos throughout gestation. Embryonic mortality strongly affects litter size and directly impacts porcine profitability. **Findings**. The expression level of 17 candidate genes selected by their known role in pregnancy establishment and development was analyze using a QuantStudio 12K Flex Real-Time PCR System in 36 endometrial samples of Iberian x Meishan F_2 sows at day 30–32 of its gestation, classified according to their estimated breeding value (EBV) as high (H, EBV>0) and low (L, EBV<0) prolificacy phenotypes. Significant differences were validated for genes: *ADM* (p=0.001; H/L ratio=3.34), *CES1* (p=0.008; H/L ratio=3.63), *FXYD3* (p=0.013; H/L ratio=2.41), *IHH* (p=0.05; H/L ratio=0.50), *KLF5* (p=0.001; H/L ratio=3.64), *KLK1* (p=0.017; H/L ratio=21.33), *PION* (p=0.009; H/L ratio=1.64), *SDCBP-2* (p=0.028; H/L ratio=2.21) and strong differences were also validated for: *DCLK-2* (p=0.07; H/L ratio=0.59), *MMP23-B* (p=0.07; H/L ratio=0.54), *NMU* (p=0.09; H/L ratio=6.81) and *SH3BGR* (p=0.09; H/L ratio=0.45). **Conclusions**. Our results have allowed the identification of new porcine genes displaying high expression differences between sows with extreme phenotypes for reproductive capacity that may play key role on the genetic architecture of prolificacy-related traits and embryo implantation failure in pigs.

Keywords: *Sus scrofa*, endometrium, gene expression, reproduction, prolificacy, litter size, OpenArray

FINDINGS

Background

Improvement in litter size has become of great interest in the pig industry as it directly impacts the productivity of the sows. Although many selection schemes have been used in order to improve this trait, its complex genetic regulation and its low heritability have made this goal very challenging (Johnson et al. 1999). There are several component traits affecting litter size: ovulation rate (OR), uterus capacity and particularly embryo survival and pre-weaning losses which can occur at each pregnancy stage an represent an average of 20‒30% (Pope & First 1985; Spötter & Distl 2006). Large phenotypic and genetic variation has been found between porcine breeds regarding litter size (Bradford 1979). The most extreme phenotypes have been observed between the Chinese Meishan, considered a very high prolific breed (Bidanel 1993) and the Iberian pig, considered one of the breeds with lowest prolificacy (Silió L; Rodriguez C; Rodrigáñez J; Toro MA 2001). Several approaches have been used to determine the factors influencing litter size. To date, more than six hundred QTLs and several candidate genes such as estrogen receptor (*ESR),* prolactin receptor *(PRLR),* follicular-stimulating hormone beta subunit (*FSHB),* erythropoietin receptor (*EPOR*), osteopontin *(OPN)* and prolactin (*PRL)* have been identified regarding this trait (Hu et al. 2013). The recent genomic revolution associated with high throughput sequencing techniques such as transcriptome analyses of different reproductive tissues have also enabled genome-wide gene expression profiling, becoming a successful strategy for identifying a higher number of candidate genes related to reproduction in livestock (Du et al. 2014; Esteve-Codina et al. 2011; Ross et al. 2009; Sun et al. 2011). We selected 18 genes predicted as differentially expressed at the transcriptome analysis previously conducted by our group (Cordoba et al. 2014) that have a high probability to play a relevant role in porcine reproduction and a positional concordance with known reproductive QTLs, to explore their expression levels in endometrial samples from pregnant sows with extreme prolificacy phenotypes (Additional file 1).

Results

The relative mRNA abundance of these 18 reproduction-related genes was analyzed in the endometrium of 36 F_2 individuals with extreme prolificacy phenotypes (H, n=18; L, n=18). Significant differences were validated for 8 genes between H and L samples (*p-value*≤0.05)

and despite not being statistically significant, strong differences were observed for genes *DCLK2*, *MMP23B, NMU* and *SH3BGR* (Figure 1). Predicted and validated expression results from this and our previous transcriptome analysis are shown at additional file 2.

Considering their function, the validated genes are mainly involved in three main stages: immune response, pregnancy establishment and embryo development and embryo implantation.

Immune response and detoxification genes: *CES1*, *SDCBP2* and *PION*

A successful embryonic implantation needs a synchronized embryo-maternal dialogue. Many components such as chemokines have an essential role in this communication leading to morphological changes during decidualization and mediating maternal acceptance towards embryo implantation (McEwan et al. 2009). Heparan sulfate proteoglycans from the syndecan (*SDC*) family such as the *SDCBP2* gene take part as co-receptors to help these chemokines bind to their innate receptors. Baston-Büst *et al.* also observed that *Sdc-1* knock-down in human endometrial cells led to dramatic changes regarding cytokine expression profiles upon decidualization and embryonic contact (Baston-Büst et al. 2010). It is possible then, that the significant increase of *SDCBP2* levels that we observe in our High prolificacy samples might support a better embryonic attachment and implantation due to the regulation of chemokine secretion of endometrial cells. Chemokines, and other pro-inflammatory factors such as cytokines are produced by different types of cells and its accumulation is triggered by the Amyloid-beta peptide leading to neuroinflammation. Formation of amyloid-beta is catalyzed by gamma-secretase activation protein (*PION*, or *GSAP*) which selectively increases its production.

In the mid-secretory phase, when the endometrium prepares for the embryo attachment, it has been observed an up-regulation of genes involved in detoxification mechanisms such as the Carboxylesterase 1, *CES1* (Giudice 2006). Carboxylesterases are a family of enzymes widely distributed among mammalian tissues with a broad range of physiologic functions, which mainly appear to hydrolyse a variety of esters, amides and thioesters. Their involvement in reproduction has been mainly described in the placenta, where are known to be involved in detoxification and metabolic activation of various drugs, environmental toxicants and carcinogens having both pharmacological and toxicological significance in the development of

the fetus. Although their role in porcine endometrium remains unknown, their detoxifying function may have a positive impact in the growing fetus, thus explaining the over expression of this gene in our high prolificacy samples.

Pregnancy establishment and embryo development genes: *ADM* and *IHH*

Adrenomedullin (*ADM*) is an hypoxia-induced vasodilator peptide highly expressed in reproductive tissues such as uterine endometrium (Hague et al. 2000), fetal membranes (Trollmann et al. 2002) and placenta(Minegishi et al. 1999). The role of this gene in fertility and implantation has been studied in several animal models. Lei *et. al.* showed that ovarian *ADM* expression appears to be involved in the regulation of progesterone production from the corpus luteum in rats (Li et al. 2011). Also in rats, Liao *et al*. pointed out the role of this gene in the regulation of embryo transport to the uterus (Liao et al. 2011). In 2008, Fritz-Six and collaborators demonstrated in mouse that homozygous deletion of *Adm* causes embryonic lethality, thus associating downregulation of *ADM* expression with several pregnancy complications (Fritz-Six et al. 2008). Its expression is regulated by several factors involved in the physiology of reproduction. In humans, it has been observed that A*DM* levels increase approximately 5-fold in the maternal plasma of normal pregnancies compared with early pregnancies specially at the earliest stages (Lenhart & Caron 2012). Our results show an over expression of this gene in high prolificacy samples, which is in agreement with these findings. Considering that one of the major differences between Iberian and Meishan prolificacy levels seems to be embryo survival, our findings suggest that this gene might be responsible of a better outcome for pregnancy establishment.

The Indian Hedgehog gene (*IHH*) it's also expressed in the uterine epithelium and its expression is progesterone-dependent (Takamoto et al. 2002). Some *in vivo* and *in vitro* studies have demonstrated that *IHH* is an essential factor that mediates the interaction between the uterine epithelium and the stroma required for achieving uterine receptivity and embryo implantation (Matsumoto et al. 2002; Takamoto et al. 2002) Unexpectedly, our results show a pronounced decrease in the expression level of this gene in high prolificacy samples. We expected that uterine under expression of this gene may lead to implantation failure caused by this deficient epithelial-stromal interaction. Although some authors have observed that a reduction in *IHH* signaling protects against bone loss (Yang et al. 2015) having a positive effect

in bone homeostasis during embryo development, from our results we cannot conclude if the role of this gene is determinant in prolificacy.

Vasculogenesis and embryo implantation genes: *KLK1* and *KLF5*

The first functional organ system to develop in the vertebrate embryo is the cardiovascular system. Embryonic growth and differentiation depends on the transport of nutrients and waste through the early vasculature (Goldie et al. 2008). Vasculogenesis is the process by which blood vessels are formed. Tissue Kallikrein gene (*KLK1*) is a potent factor with a fundamental role in vessel formation, vascular repair and robust arterializations (Stone et al. 2009). In human placenta, it has been observed that *KLK1* expression increases in first-trimester samples, suggesting that it may participate in the establishment and maintenance of placental blood flow through vasodilatation, prevention of platelet aggregation, cell proliferation, and trophoblast invasion (Valdés et al. 2001; Luo et al. 2014). Both, angiogenesis and vasculogenesis take place just a few weeks after implantation (Huppertz & Peeters 2005). We observe a highly significant decrease of *KLK1* expression in low prolificacy samples. These results suggest that defects on the expression level of these gene may underlie serious reproductive conditions, probably due to defects in the ability of trophectoderm cells to fully invade the maternal uterine wall and remodel blood vessels (Lala & Chakraborty 2003; Chaddha et al. 2004). Trophectoderm cells play key role in embryo implantation. At the beginning of the attachment reaction, the first cell type to interact with the blastocyst trophectoderm is the uterine luminal epithelium. *KLF5* gene belongs to the Kruppel-like factors (KLFs) family and its function is critical to make this uterine luminal epithelium conducive to blastocyst implantation and growth (Sun et al. 2012). In its absence, trophectoderm development is defective, resulting in developmental arrest at the blastocyst stage (Sun et al. 2012). These results suggest that *KLF5* is a key regulator of embryo pre-implantation (Lin et al. 2010) and the fact that this gene is over expressed in our high-prolificacy samples strengthens our idea of the important effect it may have on prolificacy levels and litter size control.

Conclusions

Despite prolificacy is a complex trait regulated by a intricate network of interacting genes, this work suggest a strong association between the expression level of these candidate genes and litter size control in pigs. We have provided a list of potential candidate genes that can be associated with critical steps involved in embryonic survival during the sow's gestation, thus concluding that the observed differences in the expression level of these key genes could be a determinant factor regarding prolificacy-related traits. Considering that RT-qPCR cost increases based on the number of genes being evaluated, our results also provide a set of genes on which to focus if you want to analyze differences in porcine prolificacy levels.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Experiments conceived and designed by AC, AS and SC. Experimental material collection and F_2 management: AS, CO and JN. Experiments performed by SC and AC. Data analysed by SC and AC. Statistical analysis performed by SC and AC. Paper written by SC. Critical revisions of manuscript: AC, AS, CO, IB, JN and OT. Funding obtained by AS. Research overview: AS.

REFERENCES

Baston-Büst, D.M. et al., 2010. Syndecan-1 knock-down in decidualized human endometrial stromal cells leads to significant changes in cytokine and angiogenic factor expression patterns. *Reproductive biology and endocrinology : RB&E*, 8, p.133.

Bidanel, J., 1993. Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. III. Dominance and epistatic components of heterosis on reproductive traits. *Genetics Selection Evolution*, 25(3), p.263.

Bradford, G.E., 1979. Genetic variation in prenatal survival and litter size. *Journal of animal science*, 49 Suppl 2, pp.66–74.

Chaddha, V. et al., 2004. Developmental biology of the placenta and the origins of placental insufficiency. *Seminars in fetal & neonatal medicine*, 9(5), pp.357–69.

Córdoba, S. et al., 2015. Endometrial gene expression profile of pregnant sows with extreme phenotypes for reproductive efficiency. *Scientific Reports*, 5, p.14416. Available at:

http://www.nature.com/srep/2015/151005/srep14416/full/srep14416.html [Accessed October 5, 2015].

Du, Z.-Q. et al., 2014. Identification of species-specific novel transcripts in pig reproductive tissues using RNA-seq. *Animal genetics*, 45(2), pp.198–204.

Esteve-Codina, A. et al., 2011. Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC genomics*, 12(1), p.552.

Fritz-Six, K.L. et al., 2008. Adrenomedullin signaling is necessary for murine lymphatic vascular development. *The Journal of clinical investigation*, 118(1), pp.40–50.

Giudice, L.C., 2006. Application of functional genomics to primate endometrium: insights into biological processes. *Reproductive biology and endocrinology : RB&E*, 4 Suppl 1, p.S4.

Goldie, L.C., Nix, M.K. & Hirschi, K.K., 2008. Embryonic vasculogenesis and hematopoietic specification. *Organogenesis*, 4(4), pp.257–63.

Hague, S. et al., 2000. Expression of the hypoxically regulated angiogenic factor adrenomedullin correlates with uterine leiomyoma vascular density. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 6(7), pp.2808–14.

Hu, Z.-L. et al., 2013. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic acids research*, 41(Database issue), pp.D871–9.

Hughes, P.E. & Varley, M.A., 1980. *Reproduction in the pig.* 1st ed., Massachusetts: Butterworth-Heinemann Ltd.

Huppertz, B. & Peeters, L.L.H., 2005. Vascular biology in implantation and placentation. *Angiogenesis*, 8(2), pp.157–67.

Johnson, R.K., Nielsen, M.K. & Casey, D.S., 1999. Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. *Journal of animal science*, 77(3), pp.541–57.

Lala, P.K. & Chakraborty, C., 2003. Factors regulating trophoblast migration and invasiveness: possible derangements contributing to pre-eclampsia and fetal injury. *Placenta*, 24(6), pp.575–87.

Lenhart, P.M. & Caron, K.M., 2012. Adrenomedullin and pregnancy: perspectives from animal models to humans. *Trends in Endocrinology & Metabolism*, 23(10), pp.524–532.

Li, L., O, W.-S. & Tang, F., 2011. Adrenomedullin in rat follicles and corpora lutea: expression, functions and interaction with endothelin-1. *Reproductive Biology and Endocrinology*, 9(1), p.111.

Liao, S.B. et al., 2011. Adrenomedullin increases ciliary beat frequency and decreases muscular contraction in the rat oviduct. *Reproduction (Cambridge, England)*, 141(3), pp.367–72.

Lin, S.-C.J. et al., 2010. Klf5 regulates lineage formation in the pre-implantation mouse embryo. *Development (Cambridge, England)*, 137(23), pp.3953–63.

Luo, Y. et al., 2014. Estrogen-related receptor γ serves a role in blood pressure homeostasis during pregnancy. *Molecular endocrinology (Baltimore, Md.)*, 28(6), pp.965–75.

Matsumoto, H. et al., 2002. Indian hedgehog as a progesterone-responsive factor mediating epithelial-mesenchymal interactions in the mouse uterus. *Developmental biology*, 245(2), pp.280–90.

McEwan, M. et al., 2009. Cytokine regulation during the formation of the fetal-maternal interface: focus on cell-cell adhesion and remodelling of the extra-cellular matrix. *Cytokine & growth factor reviews*, 20(3), pp.241–9.

Minegishi, T. et al., 1999. Adrenomedullin and atrial natriuretic peptide concentrations in normal pregnancy and pre-eclampsia. *Molecular human reproduction*, 5(8), pp.767–70.

Pope, W.F. & First, N.L., 1985. Factors affecting the survival of pig embryos. *Theriogenology*, 23(1), pp.91–105.

Ross, J.W. et al., 2009. Identification of differential gene expression during porcine conceptus rapid trophoblastic elongation and attachment to uterine luminal epithelium. *Physiological genomics*, 36(3), pp.140–8.

Silió L; Rodriguez C; Rodrigáñez J; Toro MA, 2001. *PORCINO IBERICO ASPECTOS CLAVES*,

Spötter, A. & Distl, O., 2006. Genetic approaches to the improvement of fertility traits in the pig. *Veterinary journal (London, England : 1997)*, 172(2), pp.234–47.

Stone, O.A. et al., 2009. Critical role of tissue kallikrein in vessel formation and maturation: implications for therapeutic revascularization. *Arteriosclerosis, thrombosis, and vascular biology*, 29(5), pp.657–64.

Sun, X. et al., 2012. Kruppel-like factor 5 (KLF5) is critical for conferring uterine receptivity to implantation. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4), pp.1145–50.

Sun, X. et al., 2011. Microarray profiling for differential gene expression in PMSG-hCG stimulated preovulatory ovarian follicles of Chinese Taihu and Large White sows. *BMC genomics*, 12, p.111.

Takamoto, N. et al., 2002. Identification of Indian hedgehog as a progesterone-responsive gene in the murine uterus. *Molecular endocrinology (Baltimore, Md.)*, 16(10), pp.2338–48.

Trollmann, R. et al., 2002. Adrenomedullin gene expression in human placental tIssue and leukocytes: a potential marker of severe tIssue hypoxia in neonates with birth asphyxia. *European journal of endocrinology / European Federation of Endocrine Societies*, 147(5), pp.711–6.

Valdés, G. et al., 2001. Tissue kallikrein in human placenta in early and late gestation. *Endocrine*, 14(2), pp.197–204.

Yang, J. et al., 2015. The Hedgehog signalling pathway in bone formation. *International journal of oral science*, 7(2), pp.73–9.

FIGURE LEGENDS

Figure 1. QuantStudio™12K Flex real-time PCR results for genes displaying significant expression differences between H and L groups. Expression values were calculated

applying the $-2^{\Delta\Delta CT}$ algorithm. Estimated relative quantities were normalized for the expression value of two uterus endogenous genes *B2MG* and *UBC* and calibrated to the sample with a higher expression. Significance was set at a *p-value* < 0.05 (*).

APPENDIX

Ethics Statement

All animal procedures were carried out according to the European animal experimentation ethics law and approved by the institutional animal ethics committee of Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Animal material and sample collection

The F_2 population used in this study comes from the cross of 3 Iberian males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) with 18 Meishan females (Domaine du Magneraud, INRA, France). The whole F_2 progeny (n=255) was obtained by matting 8 boars and 97 sows from the F_1 generation at the Nova Genètica S.A. experimental farm (Lleida, Spain). During four consecutive parities two main parameters were recorded for each sow: number of piglets born alive (NBA) and total number of piglets born (TNB) means. At day 30–32 of their fifth gestation, when litter size has reached the maximum(Hughes & Varley 1980), sows were slaughtered and the number of *corpora lutea* (CL or OR) and number of foetuses (NF) attached to the uterus were also recorded. Endometrial samples from the apical uterus of F_2 sows were collected at slaughter, and subsequently snap-frozen in liquid nitrogen. Preservation and storage was made at -80 ºC until usage.

Phenotypic records

F_2 sows were ranked into two groups: high (H; EBV>0) and low (L; EBV<0) prolificacy according to their estimated breeding value (EBV), which was calculated by using best linear unbiased predictors (BLUP) considering the reproductive traits described above: NBA and TNB means, OR and NF. Top extreme individuals from each group (H, n=18; L, n=18) were selected to be used in this study (Additional file 1).

RNA isolation and quality assessment

Total RNA was extracted using TRIzol® reagent (Invitrogen, Carlsbad, USA), following the manufacturer's instructions. The RNA integrity was assessed using an Eukaryote Total RNA Nano 6000 Labchip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA) and quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA). Only those RNA samples with an RNA integrity number (RIN) ≥ 7 were used in subsequent experiments.

Gene expression validation: QuantStudio™12K Flex Real-Time PCR System

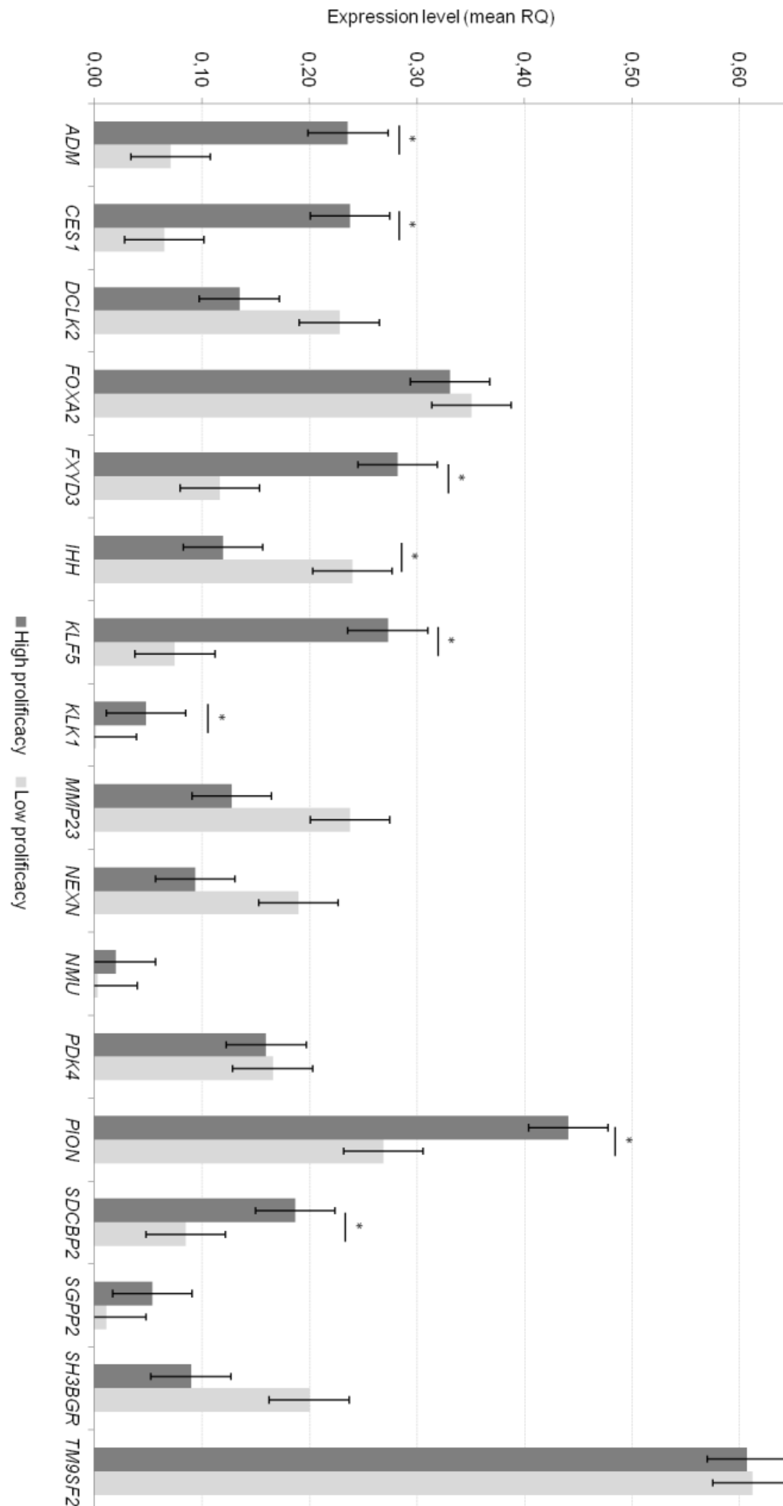
After applying these selection criteria, 18 candidate genes were selected for quantitative real-time PCR validation using a QuantStudio™ 12K Flex Real-Time PCR System in 36 extreme F_2 samples (H, n=18; L, n=18). Synthesis of cDNA for gene expression validation was performed using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems) from 1 µg of total RNA in 20 µl reaction. The synthesis of cDNA for miRNA expression validation was performed using extracted total RNA as described by Balcells *et al*. The used RT-primer sequence was 5'-CAGGTCCAGTTTTTTTTTTTTTTTVN, where V is A, C and G and N is A, C, G, and T. Minus reverse transcription (RT) and minus poly A) polymerase controls were performed. Quantitative PCR reactions were performed in triplicate in 20 µL final volume including 10 µL SYBR® Select Master Mix (Life Technologies – Thermo Fisher Scientific, Massachusetts, USA), 300 nM of each primer and 5 µL of a 1:200 dilution of the cDNA. A 1:5 relative standard curve generated from a pool of equal amounts of cDNA from all samples was included in each qPCR assay to estimate qPCR efficiency.

Reactions were incubated in a 384-well plate on a QuantStudio™ 12K Flex Real-Time PCR System. DNA primers for each gene were designed using Primer Express® software v2.0 (Applied Biosystems, USA) following manufacturer's instructions (Additional file 3).

Melting curve analysis was included in each qPCR to detect unspecific amplifications. Expression values were calculated with qbasePLUS software (Biogazelle) applying the $-2^{\Delta\Delta Ct}$ algorithm, after verifying that the assumptions of the method were met. Estimated relative quantities were calibrated to the sample with a higher expression and normalized for the

expression value of two uterus endogenous genes: *B2MG* and *UBC*. Reference genes stability was also assessed with qBasePLUS software considering a GeNorm M value < 0.5 and a coefficient of variation (CV) < 0.2. Significance was set at a *p-value*<0.05.

Figure 1.



3.2. GENE EXPRESSION REGULATION STUDIES

3.2.1. Determination of polymorphisms affecting the regulatory function of reproduction-related miRNAs

Genetic polymorphisms in miRNAs are associated with porcine extreme phenotypes for reproductive efficiency.

Córdoba-Terreros S., Balcells I., Castelló A., Ovilo C., Noguera J.L., Timoneda O., Sánchez A.

Reproduction, 2015 (under review)

Genetic polymorphisms in miRNAs are associated with porcine extreme phenotypes for reproductive efficiency.

Sarai Córdoba-Terreros[1*], Ingrid Balcells[1], Anna Castelló[1], Cristina Ovilo[2], José Luis Noguera[3], Oriol Timoneda[1], Armand Sánchez[1]

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

[2] Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Spain

[3] Àrea de Producció Animal, Institut de Recerca i Tecnologia Agroalimentàries, Universitat de Lleida, Spain

*Corresponding author:

sarai.cordoba@cragenomica.es

Edifici CRAG, Campus UAB – 08193, Bellaterra (Spain)

ABSTRACT

MicroRNAs (miRNAs) are small non-coding RNAs of 20-25 nucleotides in length that function as post-transcriptional down-regulators of gene expression. Single nucleotide polymorphisms (SNPs) affecting miRNAs have been associated with several biological processes and diseases. The aim of this study was to identify SNPs in the sequence of porcine miRNAs with a known role in reproduction to determine the effect of these variants in the reproductive efficiency of pregnant sows.

Amplified products of 9 selected miRNAs associated with reproduction traits in pigs and other mammals were Sanger-sequenced in 36 Iberian x Meishan F_2 sows classified according to their estimated breeding value (EBV), as high (H, EBV>0; n=18) and low (L, EBV<0; n=18) prolificacy. A total of 17 SNPs were identified and 13 successfully validated and genotyped in the whole intercross (n=321) by KASP™ assay. The association study performed in all F_2 samples for which EBVs were available (n=121) confirmed that variants in *ssc-mir-27a* (F=6.019; p=0.005), *ssc-mir-106a* (F=10.956, p=0.0002) and *ssc-mir-29b-2* (F=3.629, p=0.034) were significantly associated with prolificacy. Expression levels of these three mature miRNAs with a SNP significantly associated with the EBV were analyzed by RT-qPCR. We observed that the genotype for these variants has an effect on the mature miRNA expression levels: *ssc-miR-27a* (AA=0.385, AG=0.236, GG=0.739), *ssc-miR-29b-2* (AA=0.257, AG=0.016, GG=0.558) and *ssc-miR-106a* (GG=0,365. GC=0,084, CC=0.488). Our results suggest that genetic variants found in these miRNAs play a role in swine reproduction-related traits and may regulate mechanisms involved in pig litter size variation.

Keywords: *Sus scrofa*, endometrium, miRNA, polymorphisms, reproduction, prolificacy, pig, sows

INTRODUCTION

MicroRNAs are a class of small non-protein coding RNAs of approximately 20–25 nucleotides (nt) long that act mainly as post-transcriptional down-regulators of protein-coding transcripts (Bartel 2004). These small RNAs comprise one of the more abundant classes of gene regulatory molecules. MiRNAs biogenesis is temporal and spatial dependent. They are transcribed by RNA polymerase II as parts of longer primary transcripts called pri-miRNAs, which are processed to mature miRNAs in two consecutive maturation steps as described by Bartel et al.(Bartel 2004). The miRNA binds to various proteins and guides large protein complexes to partial complementary target sites, which are typically located at the 3' untranslated region (UTR) of the target mRNA (Siomi & Siomi 2010) although less common functional miRNA binding sites in the 5' UTR have recently been reported (Ørom et al. 2008). Interaction with its target gene leads to the mRNA degradation or the translational repression of its expression(Koscianska et al. 2011).

MicroRNAs processing and maturation can be directly affected by the presence of genetic variants in miRNA genes leading to a dysregulation of their expression levels. Recently, it has been demonstrated that single nucleotide polymorphisms in miRNA genes (miR-SNPs) can alter their function by modulating one or more of their processing steps (Mishra et al. 2008; Ryan et al. 2010), and as a consequence, affect the expression of several genes involved in prolificacy (L. Su et al. 2010; Wessels et al. 2013). Several authors have directly associated miR-SNPs with many relevant diseases (Zhang et al. 2015; Li et al.; Smith et al.; Hu et al.; Yang et al. 2010; Z. Hu et al. 2008), suggesting them as putative biomarkers to predict disease risk and/or prognosis (Xu & Tang 2015; Z. Hu et al. 2008). From these studies we can conclude that an aberrant miRNA expression could be directly associated with relevant biological processes in mammalian development such as embryo implantation (Chegini 2010; Pan et al. 2007; Mineno et al. 2006).

Embryo implantation and development has been widely studied in mammals for several years, especially in livestock species for its economic impact(Cha et al. 2012; R.-W. Su et al. 2010; Bazer et al. 2011). Pigs constitute economically one of the most important species in livestock but porcine profitability can be directly affected by prenatal mortality, as it constitutes a

determinant factor for litter size (Spötter & Distl 2006b). Large genetic variation has been found regarding this trait among porcine breeds and although many selection schemes have been used to improve it, its complex genetic regulation and its low heritability has made this goal very challenging(Bradford 1979; Distl 2007). To date, only a few reports have explored miRNAs expression profiles in porcine reproductive tissues, and despite miRNAs function has been related to endometrial receptivity (Sha et al. 2011; Altmäe et al. 2013; Xia et al. 2014), implantation (Chakrabarty et al. 2007; S.-J. Hu et al. 2008; Revel et al. 2011; Su et al. 2014), labor and spontaneous fetal loss in pigs (Montenegro et al. 2009; Williams, Renthal, Condon, et al. 2012; Williams, Renthal, Gerard, et al. 2012; Renthal et al. 2010; Hassan et al. 2010), miRNA-mediated regulation of sow's pregnancy remains unclear.

The goal of our study was to perform a structural and functional characterization of 9 known porcine reproduction-related miRNAs in an F_2 population coming from two porcine breeds with extreme prolificacy phenotypes, in order to determine the association of the identified structural variants with the observed differences regarding litter size, and to examine if miRNA variants could have a direct impact on their expression.

MATERIALS AND METHODS

1. Animals and sample collection

This study was performed using an F_2 population resulting from the cross of 3 Iberian (Ib) males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) with 18 Meishan (Me) females (Domaine du Magneraud, INRA, France). Once the F_1 generation was obtained, 8 boars and 97 sows were mated to obtain a 255 F_2 progeny at the Nova Genètica S.A. experimental farm (Lleida, Spain). During 4 consecutive parities, number of piglets born alive (NBA) and total number of piglets born (TNB) were recorded for 121 of these F_2 sows. At day 30-32 of the fifth gestation, when litter size potential has reached a maximum (Spötter & Distl 2006b), these sows were slaughtered and number of *corpora lutea* (OR) and number of fetuses (NF) attached to the uterus were also recorded. EBVs were then calculated by using best linear unbiased predictors (BLUP) according to these parameters: NBA and TNB means, OR and NF. At slaughter, endometrial samples from apical uterus were collected and snap frozen in liquid nitrogen. Preservation and storage was made at -80 ºC until usage. All animal procedures were

carried out according to the European animal experimentation ethics law and approved by the institutional animal ethics committee of IRTA.

2. Identification of polymorphisms

The amplified products of 9 reproduction-related miRNAs (*ssc-miR-27a*, *ssc-miR-29b-1*, *ssc-miR-29b-2*, *ssc-miR-106a*, *ssc-miR-135-1*, *ssc-miR-146a*, *ssc-miR-195*, *ssc-miR-222*, and *ssc-miR-335)* were Sanger-sequenced from endometrial genomic DNA (gDNA) in 36 extreme Ib x Me F_2 samples ranked by their EBV as high (H, EBV>0; n=18) and low (L, EBV<0; n=18) prolificacy (Supplementary Table 1). Amplified products of these 9 miRNAs were also Sanger-sequenced in 21 parental F_0 samples (Ib, n=3; Me, n=18) and 10 F_1 samples to ensure if the identified variants were segregating in our population. PCR primers design was performed on these gDNA samples to amplify the mature miRNA (miRBase v.19) ± 400 bp at the 3' and 5' respectively (Table 1). Primers were designed using Primer Express® software v2.0 (Applied Biosystems, USA) following manufacturer's instructions. All primer sequences were screened across the *Sus scrofa* genome (v. 10.2) using Primer Blast online tool to ensure their specificity.

PCRs were performed in a final volume of 25 µl containing: 5 units of AmpliTaq Gold, 2.5 mM $MgCl_2$, 1.5 mM of each dNTP, 5 µM of each primer and 2 µl of cDNA (50 ng of gDNA) for miRs *ssc-mir-27a*, *ssc-mir-135-1* and 1 unit of EcoTaq, 2 mM $MgCl_2$, 1 mM of each dNTP, 5 µM of each primer and 2 µl of cDNA for miRs *ssc-mir-29b-1*, *ssc-mir-29b-2*, *ssc-mir-106a*, *ssc-mir-146a, ssc-miR-195*, *ssc-miR-222* and *ssc-mir-335*. Reactions using AmpliTaq Gold were carried out in a thermal profile of: 95°C for 10 min, 35 cycles of 95°C for 1 min, 61°C–63°C for 1 min (depending on the primers; see table 1), 72°C for 1 min and a final extension of 72°C for 7 min. Reactions using EcoTaq were carried out in a thermal profile of: 95°C for 3 min, 35 cycles of 95°C for 1 min, 61°C–63°C for 1 min (depending on the primers; see table 1) and 72°C for 1 min, with a final extension of 72°C for 7 min. Subsequently, obtained PCR products were purified using the ExoSAP-IT® method following manufacturer's instructions. Sequencing reactions were carried out in an ABI 3730 DNA analyzer using Big Dye v1.1 terminator mixture (Applied Biosystems, Carlsbad, CA) and the same primers used for PCR

amplifications.Thermocycling profile used was: 96 °C for 1 min, 25 cycles at 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min.

3. Polymorphisms validation: KASP™ genotyping

Genomic DNA samples from the whole intercross individuals (n=321) were transferred into 96-well plates and genotyped for the identified SNPs at LGC Genomics (Herz, UK) using the KASP™ competitive allele specific PCR genotyping technology. Moreover, in order to have a preliminary idea of the putative consequences that the identified variants could have, latest version of the Variant Effect Prediction tool (VEP tool) from Ensembl was used (McLaren et al. 2010).

4. Association Study: Genotype vs. Phenotype

The association between each variant and the prolificacy phenotype was estimated using the genotype association module from the SNP & Variation Suite version 7x from Golden Helix (Golden Helix Inc., Bozeman, MT). The adjusted phenotype was fit to every encoded genotype under an additive, dominant and recessive model assumption. Student's t-test was used to estimate the association between the observed genotype for each variant and the EBV. All statistical analyses were carried out in the whole F_2 population with available EBVs (n=121). All tests were two-sided, with an $\alpha = 0.05$. Significance was set at a threshold of $p < 0.05$.

5. Variant effect on mature miRNA expression

MiRNAs displaying a SNP significantly associated with prolificacy phenotype (*ssc-miR-27a*, *ssc-miR-29b-2* and *ssc-miR-106a*) were selected to perform RT-qPCR expression validations in the same 36 extreme F_2 samples used for the SNP identification by Sanger sequencing. DNA primers (Table 2) were designed following the methodology suggested by Balcells *et al*. (Ingrid Balcells, Cirera, et al. 2011).

5.1. RNA isolation and cDNA synthesis

Endometrial total RNA was extracted from extreme F_2 sows (H, n=18; L, n=18) using TRIzol® reagent (Invitrogen, Carlsbad, USA) following manufacturer's instructions. Among them, homozygous and heterozygous samples for each identified variant were selected for RT-qPCR

validations (*ssc-mir-27a* AA=11, GG=6; *ssc-mir-29b-2* AA=16, GG=15; *ssc-mir-106a* GG=7, CC=14). RNA integrity was assessed using an Eukaryote Total RNA Nano 6000 Labchip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA) and quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA). Only those RNA samples with an RNA integrity number (RIN) ≥ 7 were used in subsequent experiments. Synthesis of cDNA was performed using total RNA as described by Balcells *et al.*(Ingrid Balcells, Cirera, et al. 2011). Minus reverse transcription (RT) and minus poly A polymerase (PAP) controls were used.

5.2. RT-qPCR

Quantitative PCR reactions were performed in triplicate in 20 µL final volume including 10 µL SYBR® Select Master Mix (Life Technologies - Thermo Fisher Scientific, Massachusetts, USA), specific primer concentrations (Table 2) and 5 µL of a 1:200 dilution of the cDNA. Relative standard curves were included in each qPCR assay to estimate target specific amplification efficiencies. Reactions were incubated in a 96-well plate at 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 sec and 60 °C for 1 min on a 7900HT Real-Time PCR System. Expression values were calculated with qbasePLUS software using target specific amplification efficiencies. Relative quantities were normalized for the expression value of *ssc-miR-103,* one of the most stable reference miRNAs in pig uterus (Timoneda et al. 2012), and calibrated to the sample with a higher expression. Reference miRNAs stability was determined considering a GeNorm M value < 0,5 and a coefficient of variation (CV) < 0,2. Significance was set at a *p-value* < 0.05.

RESULTS

1. miRNA sequencing and SNP identification

The regulatory role that some specific miRNAs exert in the expression of reproductive-related genes has been described by many authors, however, only few studies have investigated if the presence of SNPs in the sequence of these miRNAs affects their biosynthesis and function having an impact on reproduction. We have performed a functional characterization of 9 porcine miRNAs that have been reported to play a role on the regulation of relevant reproduction-related genes to identify the presence of SNPs in their sequences identifying a total of 17 putative polymorphisms. From these 17 SNPs, 13 were validated in 6 of the 9

analyzed miRNAs: *ssc-mir-27a* (n=1), *ssc-mir-29b-1* (n=2), *ssc-mir-29b-2* (n=2), *ssc-mir-106a* (n=1), *ssc-mir-135-1* (n=6) and *ssc-mir-146a* (n=1). Genomic positions and allele frequencies are shown in **table 3.** No variants were identified in *ssc-mir-195, ssc-mir-222* or *ssc-mir-335* sequences.

2. Genotyping and SNP validation

After genotyping the whole intercross (n=321) using the KASP competitive allele specific PCR genotyping technology, the 13 identified variants were validated (Supplementary Table 2). To get an initial idea of the consequences of these variants and predict their functional consequences we ran the Variant Effect Predictor (VEP) tool from Ensembl (McLaren et al. 2010). Results indicated that 47% of the variants represent an upstream gene variant, 38% represent a downstream gene variant (sequence variant located on the 3' region of a gene), 6% a non-coding transcript variant (a transcript variant of a non-coding RNA gene), 6% a non-coding transcript exon variant (a sequence variant that changes non-coding exon sequence in a non-coding transcript) and only 3% represent an intronic variant, which are transcript variants occurring within an intron.. Moreover, results indicated that our miR-SNPs positions overlap with three coding genes (*GLYCTK*, *SLU7*, *ZSWIM4*) and 15 non-coding transcripts.

3. Association study: miRNA's variant effect on prolificacy phenotype

Before the performance of the association study, data normality was assessed using the R package Shapiro-Wilk normality test: EBV (W = 0.9598, p-value = 0.001). To determine if the presence of the polymorphisms could be related with the extreme prolificacy phenotypes observed in our F_2 population, we conducted an association study. Genetic variants identified in the sequence of *ssc-mir-27a*, *ssc-mir-106a* and one of the two variants identified in s*sc-mir-29b-2* (A/G) show a significant association with sows' H or L prolificacy phenotype (Table 4). Prolificacy phenotypes were assigned based on the EBVs of the pregnant sows, thus to estimate the effect of the genotype for these three significant variants (*ssc-mir-27a* [A/G], s*sc-mir-29b-2* [A/G] and *ssc-mir-106a* [G/C]) in the EBVs of the pregnant sows, we performed a Student's t-test (Supplementary Table 3). For *ssc-mir-27a,* homozygous for the variant allele show a significant decrease on their EBV compared to homozygous for the reference allele (GG *vs*. AA; p=0.004) and although not being significant, they also present a decrease

compared to heterozygous samples (GG *vs*. AG, p=0.06). No significant differences are shown between the EBV of heterozygous samples and homozygous for the reference allele (AA *vs*. AG; p=0,424) (Figure 1a)

In *ssc-mir-29b-2*, homozygous samples for the variant present an increase on their EBV compared to the homozygous for the reference allele GG *vs*. AA; p= 0.06). Significant differences on the EBV are observed between heterozygous samples and homozygous for the reference allele (AA *vs*. AG; p=0.05). No differences regarding the EBV were found between both homozygous (AA *vs*. GG; p=0.161) or heterozygous and homozygous for the variant (AG *vs*. GG; p=0.680) (Figure 1b).

The most significant differences were found for *ssc-mir-106*. As we have observed for *ssc-mir-27a*, homozygous for the variant present a significant decrease on their EBV compared to both, homozygous for the reference allele (CC *vs*. GG; p= 0,0005) and heterozygous (CC *v*s. GC; p=0.002). However, in this particular case we observed significant differences between heterozygous and homozygous for the reference allele (GG *vs*. GC; p=0.025) (Figure 1c).

4. Mature miRNA expression validation: RT-qPCR

After assessing the putative effect that these variants which are significantly associated to our prolificacy phenotype have on the EBVs of the pregnant sows, we hypothesized if the presence of these SNPs could have an impact on the mature miRNA expression itself. RT-qPCR results indicated that all three variants identified at *ssc-mir-27a*, *ssc-mir-29b-2* and *ssc-mir-106a* precursor sequences involve an increase on the expression level of the mature miRNA, (Supplementary Table 4). In all cases, heterozygous samples present decreased expression levels of the mature miRNA, being statistically significant for both, *ssc-mir-29b-2* and *ssc-mir-106a.* For these two miRNAs, heterozygous have a significantly reduced miRNA expression levels compared to both, homozygous for the variant and for the reference allele (Figure 2a-b).

However, for *ssc-mir-27a* this decrease is statistically significant only when compared to the homozygous for the variant allele (Figure 2c).

DISCUSSION

A growing body of evidence demonstrates that aberrant miRNA expressions are associated with reproductive diseases (Pan et al. 2007; Enquobahrie et al. 2011). To date, many studies have demonstrated that these small non-coding RNAs can regulate uterine gene expression at the pre-implantation stage and also participate in placenta development and maternal-fetal interactions (Bidarimath et al. 2014). Since the function of a miRNA is mainly the repression of its target gene expression, it is clear that any alteration of the expression level of these miRNAs that regulate reproductive-related genes could be associated to an alteration of embryo implantation and development.

Polymorphisms in either the primary or precursor form of a miRNA have relevant functional implications(Z. Hu et al. 2008) and could affect mature microRNA expression either positively or negatively(Han et al. 2013). Our results demonstrate that the SNPs identified at the precursor sequences of *miR-27a* [A/G] (2:65582002), *miR-29b-2* [A/G] (9:148552571) and *miR-106a* [G/C] (X: 126200101) are significantly associated with prolificacy phenotype in our population.

Members of the miR-29 family (which include miR-29a, miR-29b-1, miR-29b-2 and miR-29c) have been proposed as potent immune gene modulators(Liston et al. 2012). In most mammals, including pigs, during early pregnancy there is an enrichment of immune cells, such as natural killer (NK) cells, T cells, B cells and macrophages at the maternal endometrium (Engelhardt 2002). These immune cells located at the maternal–fetal interface interact with foetal trophoblast cells allowing the growing foetus to develop its immunity (Erlebacher 2013). Thus, the immune response appears to play an important role in reproductive failures(Kwak-Kim et al. 2014). We have observed that the presence of the variant identified at the sequence of the *ssc-mir-29b-2* is associated with higher EBVs, being significantly higher in heterozygous individuals, compared with homozygous individuals for both, the mutant and the reference allele. These results lead us to consider not only the possibility that this particular SNP confers a better outcome in terms of litter size but the fact of being an heterozygous is associated with the best scenario in terms of prolificacy levels. Besides its role as a mediator of the immune response, *miR-29b* has been found to be involved in the inhibition of trophoblast differentiation(Kumar et al. 2013), gene reprogramming during endometrial stromal cell (ESC)

decidualization (Qian et al. 2009) and in pre-eclampsia(Li et al. 2013). Based on this and considering that RT-qPCR results show that the presence of the variant in this miRNA sequence significantly increases mature miRNA expression, we could hypothesize that having the mutant allele, would confer a better outcome in terms of immunity that will result in higher prolificacy levels (high EBVs).

In this study we have also identified a significant association between the polymorphism at the *ssc-mir-27a* and the prolificacy phenotype. Three years ago, Lei *et al.* identified another SNP in this miRNA that was significantly associated with litter size in Large White and Meishan pigs (Lei et al. 2011). In our population, we have observed a significant decrease of the EBV in homozygous samples for the variant. RT-qPCR results, however, show a significant increase of the mature miRNA expression on those homozygous samples. This leads us to hypothesize that a higher expression level of the mature miRNA results in a stronger regulation of any reproductive-related genes that would have as a consequence a decrease on the EBV.

Finally, our validations focused also on ssc-mir-106a. MiR-106a-363 family, has been found to exert an inhibition of trophoblast differentiation (Kumar et al. 2013). We have observed that the presence of the variant has a significant impact on prolificacy levels and sows with CC genotype have decreased EBVs, and thus lower prolificacy levels. RT-qPCR results show that CC genotype involves an increase on the expression level of the mature miRNA. We hypothesize that higher expression values of this mature miRNA could have a negative effect on embryo attachment because of its role in the inhibition of trophoblast. Defects in the ability of trophectoderm cells to fully invade the maternal uterine wall and remodel blood vessels has been found to lead to defective embryo implantation (Lala & Chakraborty 2003; Chaddha et al. 2004) and this may explain the observed decrease on the EBVs of homozygous for the variant.

Our results demonstrate that variants found at the precursor level may influence the biosynthesis of the mature miRNA by interfering in its biosynthesis machinery in a positive way, leading to the observed increase of its expression level. However, this may alter the regulation process that these miRNAs exert on their target genes whose expression level plays key role in mechanisms involved in pig litter size variation. Our study represents one of the firsts to explore the consequence that structural changes in miRNA precursor sequences could have on

mechanisms that mediate embryonic survival in the pig, providing the knowledge to enhance fertility and reproductive health in this species by using miR-SNPs as biomarkers.

DECLARATION OF INTEREST

The authors declare that they have no competing interests.

FUNDING

This research was funded by the Science and Innovation Ministry from Spanish Government (projects AGL2010-22358-C02-01 and AGL2004–08368-C03) and also by the Consolider-Ingenio 2010 Program (CSD2007-00036). SC is recipient of an FPI PhD fellowship from the Science and Innovation Ministry from Spanish Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Altmäe, S. et al., 2013. MicroRNAs miR-30b, miR-30d, and miR-494 regulate human endometrial receptivity. *Reproductive sciences (Thousand Oaks, Calif.)*, 20(3), pp.308–17.

Balcells, I., Cirera, S. & Busk, P.K., 2011. Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC biotechnology*, 11, p.70.

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2), pp.281–97.

Bidarimath, M. et al., 2014. MicroRNAs, immune cells and pregnancy. *Cellular & molecular immunology*, 11(6), pp.538–47.

Chaddha, V. et al., 2004. Developmental biology of the placenta and the origins of placental insufficiency. *Seminars in fetal & neonatal medicine*, 9(5), pp.357–69.

Chakrabarty, A. et al., 2007. MicroRNA regulation of cyclooxygenase-2 during embryo implantation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38), pp.15144–9.

Chegini, N., 2010. Uterine microRNA signature and consequence of their dysregulation in uterine disorders. *Animal reproduction / Colegio Brasileiro de Reproducao Animal*, 7(3), pp.117–128.

Engelhardt, H., 2002. Conceptus Influences the Distribution of Uterine Leukocytes During Early Porcine Pregnancy. *Biology of Reproduction*, 66(6), pp.1875–1880.

Enquobahrie, D.A. et al., 2011. Placental microRNA expression in pregnancies complicated by preeclampsia. *American journal of obstetrics and gynecology*, 204(2), pp.178.e12–21.

Erlebacher, A., 2013. Immunology of the maternal-fetal interface. *Annual review of immunology*, 31, pp.387–411.

Han, S.-J. et al., 2013. Kaposi's sarcoma-associated herpesvirus microRNA single-nucleotide polymorphisms identified in clinical samples can affect microRNA processing, level of expression, and silencing activity. *Journal of virology*, 87(22), pp.12237–48.

Hassan, S.S. et al., 2010. MicroRNA expression profiling of the human uterine cervix after term labor and delivery. *American Journal of Obstetrics and Gynecology*, 202(1), pp.80.e1–80.e8.

Hu, S.-J. et al., 2008. MicroRNA expression and regulation in mouse uterus during embryo implantation. *The Journal of biological chemistry*, 283(34), pp.23473–84.

Hu, Y. et al., Two common SNPs in pri-miR-125a alter the mature miRNA expression and associate with recurrent pregnancy loss in a Han-Chinese population. *RNA biology*, 8(5), pp.861–72.

Hu, Z. et al., 2008. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *The Journal of clinical investigation*, 118(7), pp.2600–8.

Koscianska, E., Starega-Roslan, J. & Krzyzosiak, W.J., 2011. The role of Dicer protein partners in the processing of microRNA precursors. *PloS one*, 6(12), p.e28548.

Kumar, P. et al., 2013. The c-Myc-regulated microRNA-17~92 (miR-17~92) and miR-106a~363 clusters target hCYP19A1 and hGCM1 to inhibit human trophoblast differentiation. *Molecular and cellular biology*, 33(9), pp.1782–96.

Kwak-Kim, J. et al., 2014. Immunological modes of pregnancy loss: inflammation, immune effectors, and stress. *American journal of reproductive immunology (New York, N.Y. : 1989)*, 72(2), pp.129–40.

Lala, P.K. & Chakraborty, C., 2003. Factors regulating trophoblast migration and invasiveness: possible derangements contributing to pre-eclampsia and fetal injury. *Placenta*, 24(6), pp.575–87.

Lei, B. et al., 2011. A SNP in the miR-27a gene is associated with litter size in pigs. *Molecular biology reports*, 38(6), pp.3725–9.

Li, C. et al., Molecular characterisation of porcine miR-155 and its regulatory roles in the TLR3/TLR4 pathways. *Developmental and comparative immunology*, 39(1-2), pp.110–6.

Li, P. et al., 2013. microRNA-29b contributes to pre-eclampsia through its effects on apoptosis, invasion and angiogenesis of trophoblast cells. *Clinical science (London, England : 1979)*, 124(1), pp.27–40.

Liston, A. et al., 2012. MicroRNA-29 in the adaptive immune system: setting the threshold. *Cellular and molecular life sciences : CMLS*, 69(21), pp.3533–41.

McLaren, W. et al., 2010. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics (Oxford, England)*, 26(16), pp.2069–70.

Mineno, J. et al., 2006. The expression profile of microRNAs in mouse embryos. *Nucleic acids research*, 34(6), pp.1765–71.

Mishra, P.J. et al., 2008. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics. *Cell cycle (Georgetown, Tex.)*, 7(7), pp.853–8.

Montenegro, D. et al., 2009. Expression patterns of microRNAs in the chorioamniotic membranes: a role for microRNAs in human pregnancy and parturition. *The Journal of pathology*, 217(1), pp.113–21.

Ørom, U.A., Nielsen, F.C. & Lund, A.H., 2008. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular cell*, 30(4), pp.460–71.

Pan, Q. et al., 2007. The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. *Molecular human reproduction*, 13(11), pp.797–806.

Qian, K. et al., 2009. Hsa-miR-222 is involved in differentiation of endometrial stromal cells in vitro. *Endocrinology*, 150(10), pp.4734–43.

Renthal, N.E. et al., 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(48), pp.20828–33.

Revel, A. et al., 2011. MicroRNAs are associated with human embryo implantation defects. *Human reproduction (Oxford, England)*, 26(10), pp.2830–40.

Ryan, B.M., Robles, A.I. & Harris, C.C., 2010. Genetic variation in microRNA networks: the implications for cancer research. *Nature reviews. Cancer*, 10(6), pp.389–402.

Sha, A.-G. et al., 2011. Genome-wide identification of micro-ribonucleic acids associated with human endometrial receptivity in natural and stimulated cycles by deep sequencing. *Fertility and sterility*, 96(1), pp.150–155.e5.

Siomi, H. & Siomi, M.C., 2010. Posttranscriptional regulation of microRNA biogenesis in animals. *Molecular cell*, 38(3), pp.323–32.

Smith, R.A. et al., A genetic variant located in miR-423 is associated with reduced breast cancer risk. *Cancer genomics & proteomics*, 9(3), pp.115–8.

Spötter, A. & Distl, O., 2006. Genetic approaches to the improvement of fertility traits in the pig. *Veterinary journal (London, England : 1997)*, 172(2), pp.234–47.

Su, L. et al., 2010. Differential expression of microRNAs in porcine placentas on days 30 and 90 of gestation. *Reproduction, fertility, and development*, 22(8), pp.1175–82.

Su, L. et al., 2014. Expression patterns of microRNAs in porcine endometrium and their potential roles in embryo implantation and placentation. *PloS one*, 9(2), p.e87867.

Timoneda, O. et al., 2012. Determination of reference microRNAs for relative quantification in porcine tissues. *PloS one*, 7(9), p.e44413.

Wessels, J.M. et al., 2013. The microRNAome of pregnancy: deciphering miRNA networks at the maternal-fetal interface. *PloS one*, 8(11), p.e72264.

Williams, K.C., Renthal, N.E., Condon, J.C., et al., 2012. MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), pp.7529–34.

Williams, K.C., Renthal, N.E., Gerard, R.D., et al., 2012. The microRNA (miR)-199a/214 cluster mediates opposing effects of progesterone and estrogen on uterine contractility during pregnancy and labor. *Molecular endocrinology (Baltimore, Md.)*, 26(11), pp.1857–67.

Xia, H.-F. et al., 2014. MicroRNA expression and regulation in the uterus during embryo implantation in rat. *The FEBS journal*, 281(7), pp.1872–91.

Xu, L. & Tang, W., 2015. The associations of nucleotide polymorphisms in mir-196a2, mir-146a, mir-149 with lung cancer risk. *Cancer biomarkers : section A of Disease markers*, 15(1), pp.57–63.

Yang, R. et al., 2010. A genetic variant in the pre-miR-27a oncogene is associated with a reduced familial breast cancer risk. *Breast cancer research and treatment*, 121(3), pp.693–702.

Zhang, X. et al., 2015. Association of the miR-146a , miR-149 , miR-196a2 and miR-499 polymorphisms with susceptibility to pulmonary tuberculosis in the Chinese Uygur, Kazak and Southern Han populations. *BMC infectious diseases*, 15(1), p.41.

FIGURE LEGENDS

Figure 1 (a-c). Association between the genotype for the identified variants and the EBVs of the pregnant sows. Student's t-test was used to estimate the association between the observed genotype for each variant and the EBVs. Statistical analysis was carried out in the F_2 samples for whom EBVs were available (n=121). All tests were two-sided, with a $\alpha = 0.05$. Significance was set at a threshold of $p < 0.05$.

Figure 2 (a-c). Mature miRNA expression results obtained by RT-qPCR. Relative quantities were calculated using target-specific amplification efficiencies and normalized for the expression level of the uterus reference miRNA *ssc-miR-103* (M=0.464; CV=0.166).

TABLES

Table 1. Primers used for the SNP identification by Sanger sequencing.

miRNA	Forward primer	Reverse primer	Conc.	Temp
ssc-mir-27a	CCCCAGTGGTAGGATAGGC	TCATTACCTCCTTTTGTCTCTCC	300 nM	63ºC
ssc-mir-29b-1	TGTGTACGTGGGAGATACGCT	GGGTGGTACGGATCCACTG	200 nM	61ºC
ssc-mir-29b-2	GTTGCCTGCGTACAGCTT	TCCCTTCTTGAACCGGC	200 nM	61ºC
ssc-mir-106a	CACTTTGGTACTGCCGGGAC	TGTGAGGACGGAGCAGAAGA	200 nM	63ºC
ssc-mir-135-1	GCCAGGACAGAAGGAAAGGA	CCTTTGCTAAGTGTCCCAGC	300 nM	63ºC
ssc-mir-146a	TCACATGAGTGTCAGGACTAGAC	ATAACAGCATGGAAAGCACTTA	200 nM	57ºC
ssc-mir-195	GCCTTCGTTGCCCACAC	TGCTGTTCCTGTATGAGCATC	200 nM	61ºC
ssc-mir-222	AGCTTTCACTACTGAGGACTTCC	TGCATCTGTACATGGGCTT	200 nM	57ºC
ssc-mir-335	CCAACACATATTGAAGATTTCCT	AAACGAGCTTGGAAAAGATTT	200 nM	61ºC

Table 2. Primers used for the mature miRNAs expression validation by RT-qPCR.

miRNA	Primers	Primer sequence	Conc.
ssc-miR-27a	Forward	TCGTGTTCACAGTGGCTAAGTTC	250nM
	Reverse	TCCAGTTTTTTTTTTTTTTTGCG	
ssc-miR-29b-2	Forward	CATCTTTGTATCTAGCACCATTTGAAAT	300nM
	Reverse	GGTCCAGTTTTTTTTTTTTTTTAACACT	
ssc-miR-106a	Forward	CGTGTAAAAGTGCTTACAGTGCAG	500nM
	Reverse	GTCCAGTTTTTTTTTTTTTTTGCTAC	
ssc-miR-103	Forward	AGAGCAGCATTGTACAGG	250nM
	Reverse	GGTCCAGTTTTTTTTTTTTTTTCATAG	

Table 3. Identified variants by Sanger-sequencing in the sequence of the candidate microRNAs.

miRNA	SNP	Position in miRNA seq.	Coordinates (Ssc.10.2)	Observed Alleles		Major Allele (MJA)	MJA Freq.	Minor Allele (MIA)	MIA Freq.
				Iberian	Meishan				
ssc-miR-27a	[A/G]	pri-mir-27a	2:65582002	A	A/G	A	0.602	G	0.398
ssc-miR-29b-1	[T/C]	pre-mir-29b-1	18:19034822	T	T/C	T	0.948	C	0.052
	[A/G]		18:19034985	A	A/G	A	0.949	G	0.051
ssc-miR-29b-2	[G/T]	pre-mir-29b-2	9:148552568	G	G/T	G	0.76	T	0.24
	[A/G]		9:148552571	A	G	A	0.693	G	0.307
ssc-miR-106a	[G/C]	pre-mir-106a	X:126200101	G	G/C	G	0.673	C	0.327
ssc-miR-135-1	[C/T]	pre-mir-135-1	13:37563816	C	T	C	0.551	T	0.449
	[T/C]		13:37563874	T	C				
	[G/C]		13:37563891	G	C	G	0.574	C	0.426
	[A/G]		13:37563928	A	G	A	0.551	G	0.449
	[G/A]		13:37563973	G	A	G	0.602	A	0.398
	[C/T]		13:37564023	C	T	C	0.551	T	0.449
ssc-miR-146a	[A/G]	pre-mir-146a	16:68488282	A	A/G	A	0.633	G	0.367

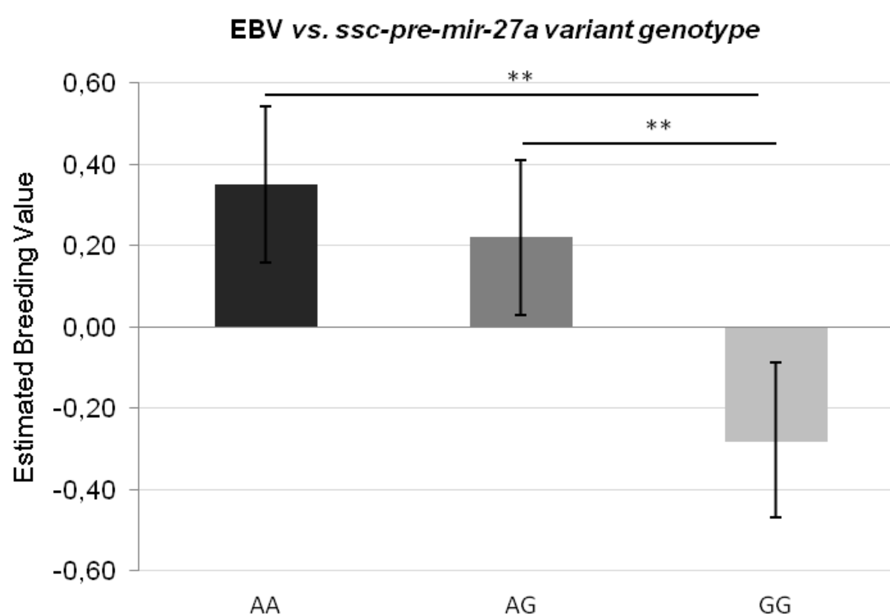
Table 4. Association test results. miRNAs in bold are those which identified variant shows a significant effect on sample's prolificacy phenotype (EBV).

miRNA	SNP	Position in miRNA seq.	Coordinates (Ssc.10.2)	Observed Alleles		Genotype *vs.* Phenotype	
				Iberian	Meishan	F	*p-value*
ssc-mir-27a	**[A/G]**	**pri-mir-27a**	**2:65582002**	**A**	**A/G**	**6,02**	**0,005**
ssc-mir-29b-1	[T/C]	pre-mir-29b-1	18:19034822	T	T/C	0,52	0,487
	[A/G]		18:19034985	A	A/G	0,06	0,809
ssc-mir-29b-2	[G/T]	**pre-mir-29b-2**	9:148552568	G	G/T	0,09	0,913
	[A/G]		**9:148552571**	**A**	**G**	**3,63**	**0,034**
ssc-mir-106a	**[G/C]**	**pre-mir-106a**	**X:126200101**	**G**	**G/C**	**10,96**	**0,000**
ssc-mir-135-1	[C/T]	pre-mir-135-1	13:37563816	C	T	0,07	0,934
	[T/C]		13:37563874	T	C	0,09	0,91
	[G/C]		13:37563891	G	C	0,08	0,92
	[A/G]		13:37563928	A	G	0,06	0,941
	[G/A]		13:37563973	G	A	0,58	0,562
	[C/T]		13:37564023	C	T	0,06	0,941
ssc-mir-146a	[A/G]	Intron of *SLU7*	16:68488282	A	A/G	1,56	0,22

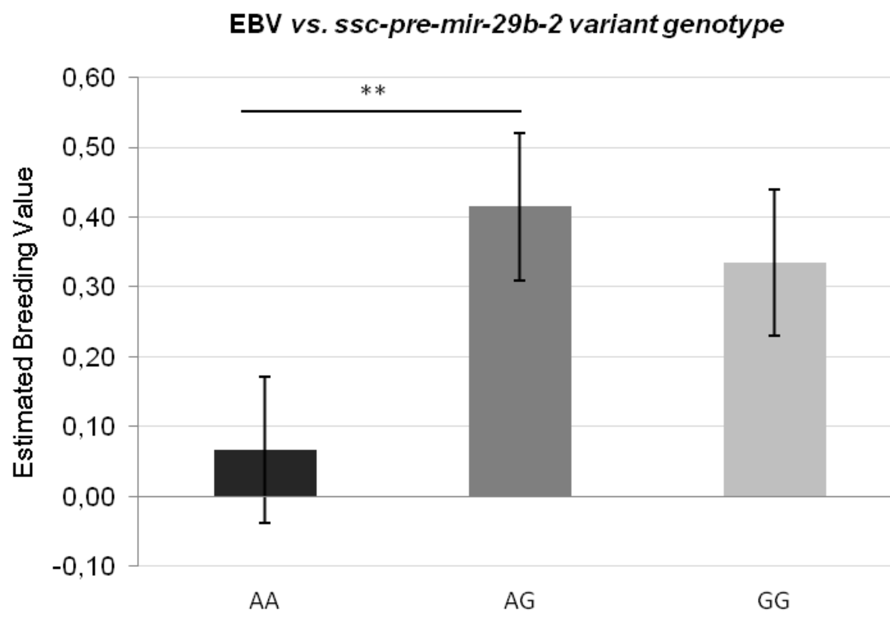
FIGURES

Figure 1.

A



B



C

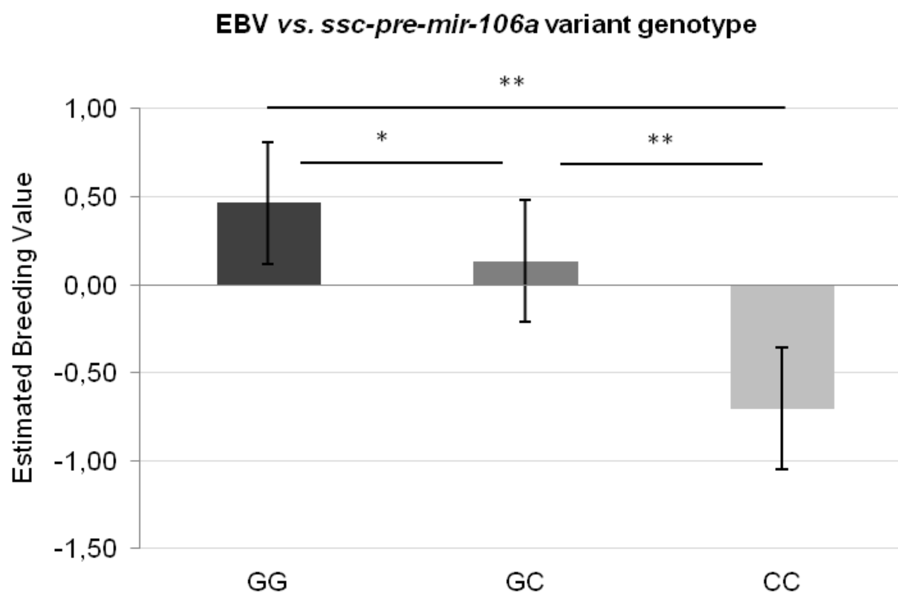
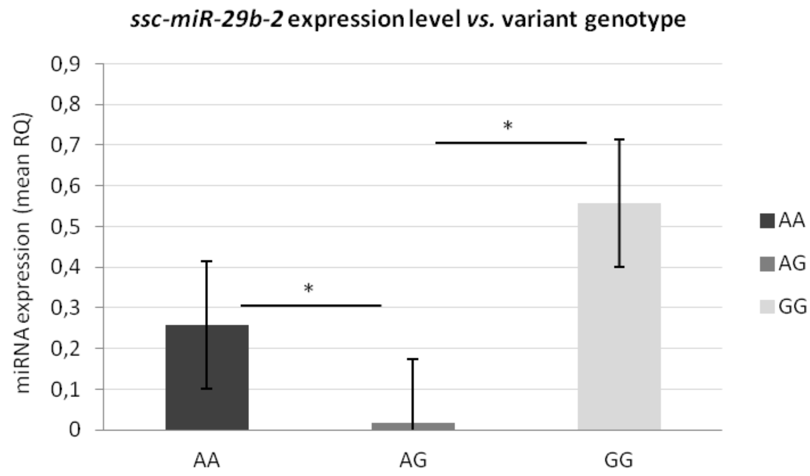
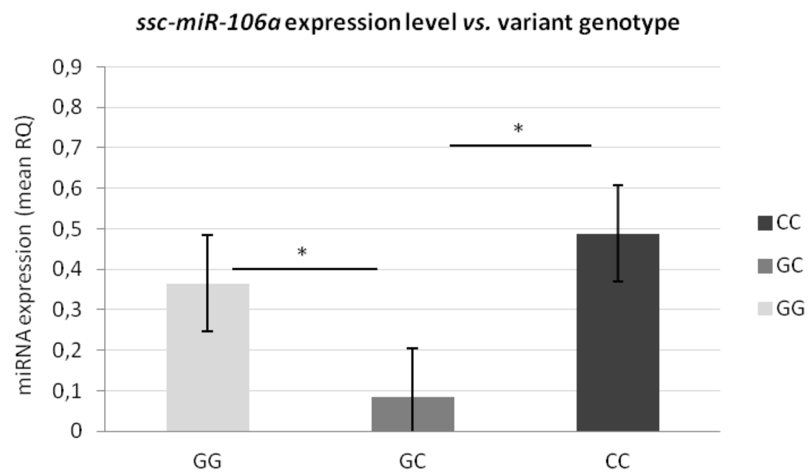


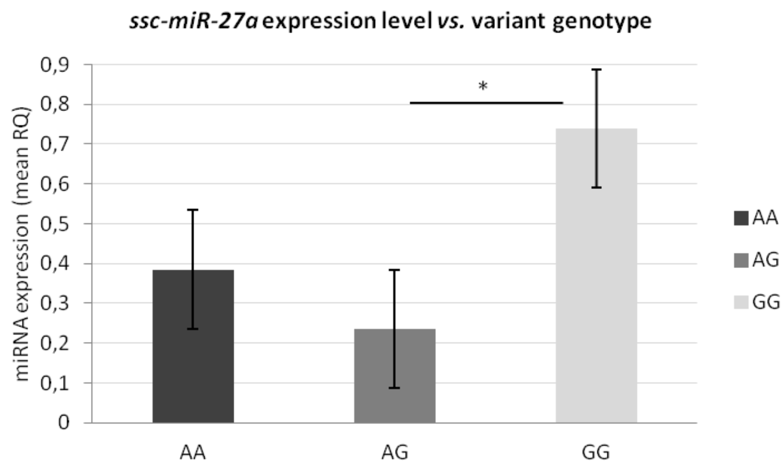
Figure 2.

A



B





3.2.2. Functional validation of the miRNA-mRNA interaction

Expression of candidate genes affecting porcine reproductive efficiency is regulated by microRNAs ssc-*miR-101-3p*, *ssc-miR-144-3p* and *ssc-miR-195-5p*

S. Córdoba[1*], S. Dhorne-Pollet[4], A. Castelló[1], I. Balcells[1], C. Ovilo[2],

JL. Noguera[3], O. Timoneda[1], A. Sánchez[1]

(Manuscript in preparation)

Expression of candidate genes affecting porcine reproductive efficiency is regulated by microRNAs ssc-*miR-101-3p, ssc-miR-144-3p* and *ssc-miR-195-5p*

S. Córdoba[1*], S. Dhorne-Pollet[4], A. Castelló[1], I. Balcells[1], C. Ovilo[2], JL. Noguera[3], O. Timoneda[1], A. Sánchez[1]

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

[2] Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (SGIT-INIA), Madrid, 28040, Spain

[3] Genètica i Millora Animal, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Lleida, 25198, Spain

[4] Animal Genetics and Integrative Biology, Institut National de la Recherche Agronomique (INRA), Jouy-en-Josas, 78354, France

*Corresponding author:

sarai.cordoba@cragenomica.es

034935636600

INTRODUCTION

MicroRNAs (miRNAs) are endogenous small non-protein coding RNAs of approximately 20–25 nucleotides long that play important regulatory roles of protein-coding transcripts in animals and plants (Bartel 2004). Their processing and maturation is temporal and spatial dependent. As described by Bartel et al. (Bartel 2004), miRNAs are first transcribed as parts of longer primary transcripts called pri-miRNAs by RNA polymerase II. Subsequently, they are processed to mature miRNAs in two consecutive maturation steps. These small non-coding regulators, bind to partial complementary target sites typically located at the 3' untranslated region (UTR) of their target mRNA (Siomi & Siomi 2010) causing either, its complete degradation or the translational repression of its expression (Koscianska et al. 2011).

Recently, miRNAs have emerged as new players in the required fine tuning of embryo development and implantation in humans and other species (Xia et al. 2014; Mineno et al. 2006; Estella et al. 2012). Embryo implantation that has been widely studied in mammals through several years, takes special relevance in livestock species for its economic impact. Pigs are one of the most important species in livestock and the main goal of porcine industry is improving reproductive efficiency since it directly improves production (Spötter & Distl 2006a). However, large genetic variation regarding prolificacy levels has been found among porcine breeds (Rothschild 1996). The complex genetic regulation of this trait and its low heritability has made the selection of this character rather challenging and although recent transcriptomic analyses have explored miRNAs expression profiles in porcine reproductive tissues, only a few studies have attempted to functionally validate miRNA-mediated regulation of reproduction-related genes and litter size control (Jin et al. 2013; Gould & Subramani 1988; Su et al. 2014). Although to date, some studies have related miRNAs with endometrial receptivity (Sha et al. 2011; Altmäe et al. 2013; Xia et al. 2014), implantation (Chakrabarty et al. 2007; S.-J. Hu et al. 2008; Revel et al. 2011; Su et al. 2014), labor and spontaneous fetal loss in pigs (Montenegro et al. 2009; Williams, Renthal, Condon, et al. 2012; Williams, Renthal, Gerard, et al. 2012; Renthal et al. 2010; Hassan et al. 2010), only a few studies have attempted to validate how these specific miRNA:mRNA interactions could affect porcine reproduction (Su et al. 2014; Lin et al. 2012; Hou et al. 2012; Xu et al. 2011).

The goal of our study was to perform a functional validation of the miRNA:mRNA interaction between four miRNAs predicted to target four porcine reproduction-related genes that have been associated with key processes involved in litter size control.

MATERIAL AND METHODS

Ethics Statement

All animal procedures were carried out according to the European animal experimentation ethics law and approved by the institutional animal ethics committee of Institut de Recerca i Tecnologia Agroalimentàries (IRTA).

Animal material and sample collection

Pregnant sows used in this study come from an F_2 resulting of the cross of 3 Iberian males from the Guadyerbas line (Dehesón del Encinar, Toledo, Spain) with 18 Meishan females (Domaine du Magneraud, INRA, France). The whole F_2 progeny (n=255) was obtained by matting 8 boars and 97 sows from the F_1 generation at the Nova Genètica S.A. experimental farm (Lleida, Spain). During four consecutive parities number of piglets born alive (NBA) and total number of piglets born (TNB) means were recorded for each sow. At day 30–32 of their fifth gestation, when litter size has reached the maximum (Hughes & Varley 1980), sows were slaughtered and the number of *corpora lutea* (CL or OR) and number of foetuses (NF) attached to the uterus were also recorded. At slaughter, endometrial samples from the apical uterus were collected and snap-frozen in liquid nitrogen. Preservation and storage was made at -80 ºC until usage.

Phenotypic records

F_2 sows were ranked into two groups: high (H; EBV>0) and low (L; EBV<0) prolificacy according to their estimated breeding value (EBV), which was calculated by using best linear unbiased predictors (BLUP) considering the reproductive traits described above (Fernandez-Rodriguez et al. 2011): NBA and TNB means, OR and NF. Top extreme individuals from each group were selected to be used in this study (Table 1).

Table 1. Phenotypic records of the F_2 Iberian × Meishan sows used in this study. [a]NBA (number of piglets born alive) and TNB (total number of piglets born) trait entries correspond to the average for four consecutive parities. [b]OR (number of *corpora lutea*) and NF (number of foetuses) recorded at slaughter on the fifth gestation.

Prolificacy level	Animal	NBA[a]	TNB[a]	OR[b]	NF[b]	EBV
HIGH	A1 (787)	11.75	13	16	16	1.68
	A2 (332)	12.75	13.33	16	14	1.55
	A3 (373)	11.25	11	20	17	1.5
Average (HIGH)		**11.92**	**12.44**	**17.33**	**15.67**	**1.58**
LOW	A4 (350)	4.5	3	15	6	-2.48
	A5 (360)	5	5.33	18	1	-2.33
	A6 (861)	5.5	5	24	9	-2.04
Average (LOW)		**5.00**	**4.44**	**19.00**	**5.33**	**-2.28**

RNA isolation and quality assessment

Total RNA was extracted using TRIzol® reagent (Invitrogen, Carlsbad, USA), following the manufacturer's instructions. The RNA integrity was assessed using an Eukaryote Total RNA Nano 6000 Labchip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, USA) and quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA). Only those RNA samples with an RNA integrity number (RIN) ≥ 7 were used in subsequent experiments.

Selection of genes and target miRNAs identificaction

Considering previous results from a transcriptome analysis in the same F_2 sows (Córdoba S. et al, 2015), we selected 4 genes predicted as differentially expressed in the RNA-seq analysis (q-value<0.05), displaying expression differences above 2 fold, and with a positive mapping into known porcine reproductive QTLs which also present significant functional enrichment in relevant pathways associated to porcine reproduction (Table 2).

Table 2. RNA-seq results for the selected genes. All expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads – Mortazavi et al., 2008). Mean difference between both groups is represented as the $\log_2$ transformed fold change (Log_2FC). False discovery rate (FDR) corrected *p-values* are represented *as q-value.*

Gene	H prolificacy (RPKM)	L prolificacy (RPKM)	Log_2 FC	p-value	q-value[a]	Up-regulation
ADM	165.93	35.82	-2.2	0.0004	0.032	High
HTRA3	19.97	77.68	2.0	0.0004	0.032	Low
PTHLH	286.10	14.36	-4.3	0.0001	0.008	High
VEGFA	992.55	60.23	-4.0	0.0001	0.008	High
ssc-miR-144	10.64	0.76	-3.8	0.0110	0.985	High
ssc-miR-101	430.21	187.19	-1.2	0.0340	0.985	High
ssc-miR-181d-5p	55.51	124.35	1.2	0.0460	0.985	Low
ssc-miR-195	2033.69	2154.79	0.1	0.8420	1.000	Low

Putative target miRNAs for these genes were computationally predicted using TargetScan 6.0 and miRDB 5.0 algorithms. As TargetScan does not host predicted gene miRNA targets in Sus scrofa, we first checked the conservation of the predicted binding sites and seed regions. Predicted miRNAs were considered true positives only if 8mer and 7mer sites were conserved and have a score > 50. MiRDB 5.0 was run in Custom prediction mode, which allowed us a more flexible miRNA target search by providing our own *Sus scrofa* 3'UTR sequences. Both softwares were run using default parameters. Among all predicted miRNAs, we selected only those for which we had previous endometrial expression evidences by RNA-seq. Only those miRNA:mRNA interactions predicted by both softwares and displaying overlapping binding sites were retained.

Gene-miRNA interaction: Luciferase Reporter Assay

Cell culture and 3'-UTRs cloning

Cell culture was performed from Human embryonic kidney 293T cells (HEK293T) since they are considered one of the most stable reporter cell lines and present several favorable properties such their origin, adhesion and growth characteristics (Agarwal et al. 2015; Yuan et al. 2015; Zhao et al. 2014). Cells were grown in DMEM supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-Glutamine, 100U/mL penicillin and 100µg/mL streptomycin.

The 3'-UTR of the four mRNA predicted to be targeted by microRNA were amplified by PCR. The amplified 3'-UTR were cloned into the cloning site downstream of firefly luciferase in the pmiRGLO dual-luciferase vector (Promega). Primers sequences and restriction enzyme used for cloning of the 3'-UTRs of porcine genes are described in Table 3. *E. coli* cells transformed with recombinant pMIRGLO Dual-Luciferase miRNA Target Expression vector were grown overnight in appropriate volume of LB medium with ampicillin (100µg/mL). Plasmid DNA was isolated using the 'Pure Yield Plasmid Miniprep system' according the manufacturer's protocol (Promega).

Table 3. Primers and restriction enzyme sequences used for the 3'-UTRs cloning.

Target	Primers	Primer sequence	Amplicon size	microRNA
HTRA3	Forward	GTTGAGCTCCCCCAAAAGGCCATCTCG	302 bp	*ssc-miR-101-3p*
	Reverse	GTTCTCGAGCGTGTGCTTGTAAACTTTAATTTCCA		
VEGFA	Forward	GTTGAGCTCGAGCCTCCCTCAGCGTTTT	507 bp	*ssc-miR-195-5p*
	Reverse	GTTCTCGAGGGATCTGGGTAGGGACGTTCTC		
ADM	Forward	GTTGAGCTCGGAGGCAGTGTTCTCTTCGG	527 bp	*ssc-miR-181d-5p*
	Reverse	GTTCTCGAGTGGTGTTTTCTTCTTCCCCAA		
PTHLH	Forward	GTTGAGCTCTTCAGAGGACGTATTGCAGAATTC	401 bp	*ssc-miR-144*
	Reverse	GTTCTCGAGATACTGCTATTTTTACATGCACAGAGG		

Transfection and Luciferase activity measurement

One day prior transfection, HEK-293T cells were seeded in 24-well plates at density 2.105 cells/well in DMEM containing 10% FBS without antibiotics. Co-transfections of each miRNA mimics (Table 1) and reporter plasmid (Promega) were performed using Dharmafect I transfection reagent (Dharmacon). Two different concentrations of miRNA inhibitors (33nM and 100nM) and two different concentrations of plasmid (50ng and 100ng) were used.
After 24h, cells were washed twice in PBS and lysed with 100µl of passive lysis buffer (Promega). An aliquot of 20µl were assayed for firefly and Renilla luciferase activity using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's protocol. Luciferase activity values were obtained using an infinite 200 Pro luminometer (Tecan). For

each putative target, control experiments were performed including a plasmid which does not contain the 3'-UTR fragment, and a scrambled microRNA.

Statistics

The Firefly luciferase activity was normalized to the Renilla luciferase activity, and then this ratio was normalized to the control constructs used in each experiment. Four independent transfections for each condition were averaged and two-tailed Student's t-tests were used to compare samples.

RESULTS

Target miRNAs identification

After running both algorithms, TargetsScan and miRDB, only those microRNAs predicted to target our candidate genes by both softwares and displaying overlapping binding sites were considered. Moreover, to enseure the success of the functional validation, only highly reliable interactions were selected by checking the Target and Context ++ scores (Table 4). After applying these criterions, 4 interactions were retained: *ADM-ssc-miR-181d-5p*, *HTRA3-ssc-miR-101-3p*, *PTHLH-ssc-miR-144-3p* and *VEGFA-ssc-miR-195-5p*.

Table 4. MiRNA targeting prediction performed by TargetScan and miRDB.

Gene	Predicted target miRNA	Conservation		TargetScan prediction			miRDB prediction	
		miRNA sequence	3'UTR seed region	Site type[a]	Context++ score percentile[b]	PC[c]	Target rank[d]	Target score[e]
VEGFA	hsa-miR-195-5p	UAGCAGCACAGAAAUAUUGGC	TGCTGCT	8mer	99	0.82	5	99
	ssc-miR-195-5p	UAGCAGCACAGAAAUAUUGGC						
PTHLH	hsa-miR-144-3p	UACAGUAUAGAUGAUGUACU	ATACTGT	8mer	99	0.25	22	89
	ssc-miR-144-3p	UACAGUAUAGAUGAUGUAC						
HTRA3	hsa-miR-101-3p	UACAGUACUGUGAUAACUGAA	GTACTGT	8mer	99	0.80	1	95
	ssc-miR-101-3p	UACAGUACUGUGAUAACUGAA						
ADM	hsa-miR-181d-5p	AACAUUCAUUGUUGUCGGUGGGU	TGAATGT	7mer-8mer	49	0.69	26	74
	ssc-miR-181d-5p	AACAUUCAUUGUUGUCGGUGGGUU						

[a]An exact match to positions 2-8 of the mature miRNA. [b]The context++ score percentile rank is the percentage of sites for this miRNA with a less favorable context++ score. [c]Probability of conserved targeting as described in Friedman et al., 2009.[d]Position that this prediction has among the whole set of predicted targets. [e]Target score is the confidence that miRDB algorithm gives to the prediction, where 100 represents the most and 50 the less likely to be real.

Gene-miRNA interaction: Luciferase Reporter Assay

Luciferase reporter assays were performed in HEK293T cells. After cloning and testing the interaction of these four reproduction-related genes with their predicted target miRNAs we observed a down regulation of 3 of these genes upon their respective target miRNAs.
For *VEGFA* gene, we detected a downregulation of 17% upon mimics with *ssc-miR-195-5p* when using 33nM of miRNA mimics and 50ng of the reporter plasmid and 42% when using 100nM of miRNA mimics and 100ng of reporter plasmid (Figure 1a). Our previous RNA-seq results showed increased levels of *VEGFA* and decreased levels of its target *ssc-miR.195-5p* in uterine endometrium of high prolificacy sows (Figure 2a). The expression level of *HTRA3* was reduced in a *34%* upon 100nM of *ssc-miR-101-3p* mimics and 100ng of the reporter plasmid (Figure 1c). In this case, previous expression results from endometrial transcriptome analysis show an increase of *HTRA3* in low prolificacy samples, which correlates with a decrease in *ssc-miR-101-3p* (Figure 2c). The highest downregulation was observed for *PTHLH* gene upon *ssc-miR-144-3p* mimics. In this case, expression levels were reduced in a 48% when using 33nM of miRNA mimics and 50ng of the reporter plasmid and in a 69% when using 100nM of miRNA mimics and 100ng of the reporter plasmid (Figure 1b). In this particular case, RNA-seq results showed a positive correlation between the expression levels of both, *PTHLH* and its target *ssc-miR-144-3p* (Figure 2b).Contrary to what previously observed in our RNA-seq data (Figure 2d), no down-regulation of *ADM* was observed in the presence of *ssc-miR-181d-5p* mimics (Figure 1d).

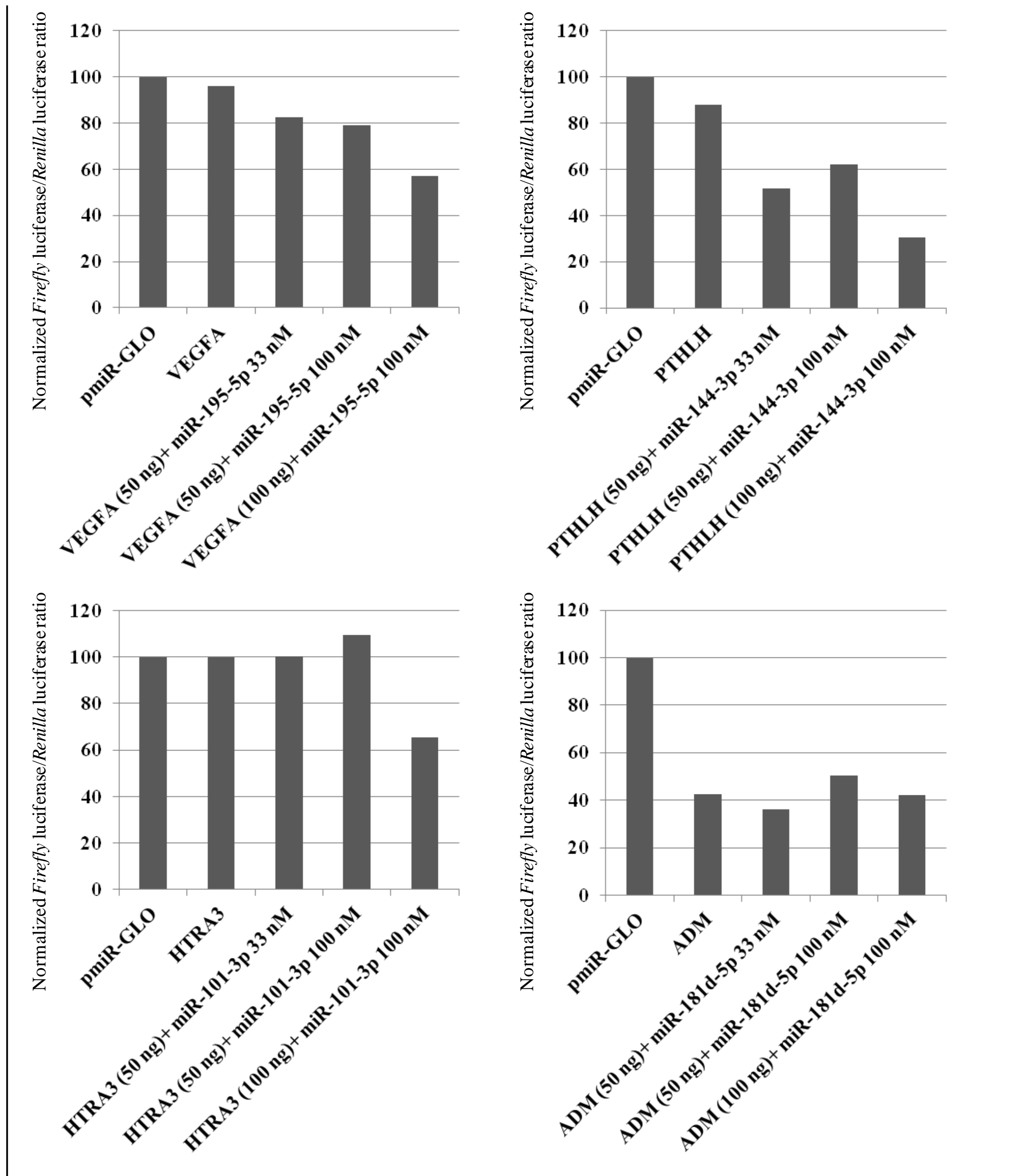


Figure 1(a-d). Luciferase reporter assay results. Firefly luciferase activity was measured and normalized by the Renilla luciferase activity. Data are represented as mean ± SEM from four independent transfection experiments. Two-tailed Student's t-tests were used to compare samples and significance was set at a $p<0.05$.

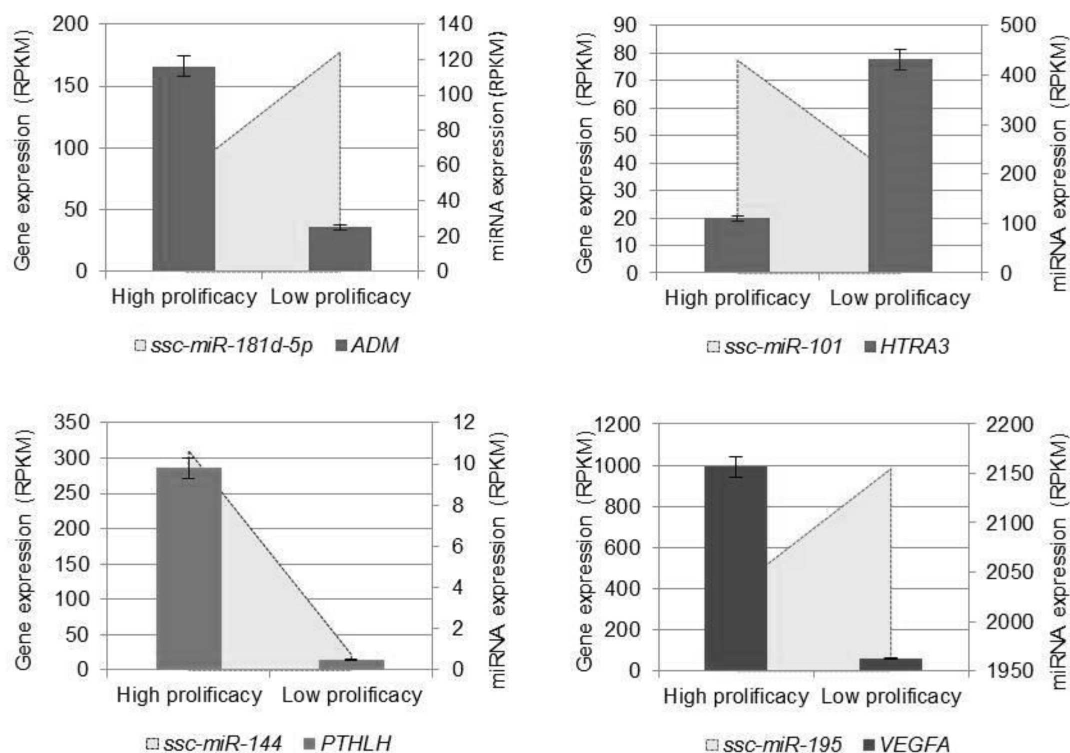


Figure 2 (a-d). Summary of the previous endometrial transcriptome analysis results.

Expression results are shown as Reads per Kilobase Mapped reads (RPKM).

DISCUSSION

MiRNAs have been widely associated with mammalian development (Tang et al. 2007). These small RNAs have key functions in many relevant biological pathways involved in embryo formation, implantation and early development. However, the exact role of miRNAs in normal embryo formation and endometrial preparation for pregnancy still remains unknown.

Pregnancy is a complex physiological process that requires fine-tuning of many factors such as hormones, growth factors and cytokines between the mother and the developing conceptus, in order to ensure a successful outcome. Any disturbance in this fine-tuning will lead to pregnant losses, which in pigs can represent up to a 45%. In a previous transcriptome analysis performed by our group, more than a hundred genes were observed as differentially expressed between the endometrium of pregnant sows with divergent prolificacy phenotypes (Córdoba et al. 2015). Some of these genes have been widely studied in several species including humans, monkeys and mice. Because of its predicted and in some cases already validated role in

reproduction, we focused our miRNA-targeting study in *ADM*, *HTRA3*, *PTHLH* and *VEGFA* genes.

Vascular endothelial growth factor A (*VEGFA*) is a member of the *PDGF/VEGF* growth factor family that has several functions, including mediating increased vascular permeability, inducing angiogenesis, vasculogenesis and endothelial cell growth, promoting cell migration, and inhibiting apoptosis. Because of its relevant role in embryonic implantation, *VEGFA* has been proposed as a candidate gene for litter size in pigs (Chen et al. 2014; Krawczynski et al. 2014). Furthermore, single nucleotide polymorphism identified in *VEGFA* was shown to be associated with this trait also in pigs. In this study, we have observed a downregulation of this gene upon *ssc-miR-195-5p* mimics that causes a decrease on its expression level of a 42%. Considering the relevant role that *VEGFA* has on litter size and reproductive pathways affecting embryo implantation, the repression that *ssc-miR-195-5p* exerts on its expression would be a clear evidence of a possible unfavorable outcome regarding prolificacy levels. Our previous RNA-seq results are in agreement with this, showing increased levels of *VEGFA* and decreased levels of its target *ssc-miR.195-5p* in uterine endometrium of high prolificacy sows. Altogether, confirms our hypothesis that the repression that this miRNA causes on its target gene may play a key role in sows' reproductive efficiency.

Another relevant gene is the HtrA (high-temperature requirement factor A) serine peptidase 3 (*HTRA3*). Although the role of this gene in porcine reproduction has not yet been elucidated, in humans, it has been reported to inhibit *TGF-β* signaling in the endometrium and has been proposed as an early marker for preeclampsia because it negatively regulates trophoblast invasion during placentation. Physical interaction between trophoblast and uterine decidual cells is a required process to ensure uterine receptivity; therefore, an overexpression of this gene may cause a defective endometrial preparation to embryo attachment. Our results have confirmed a positive interaction between *HTRA3* and its target *ssc-miR-101-3p*. In this case, gene expression decreased a 34% upon miRNA mimics. Previous expression results from endometrial transcriptome analysis show an increase of *HTRA3* in low prolificacy samples, which correlates with a decrease in *ssc-miR-101-3p*. Considering these observations in addition to luciferase results, we hypothesize that the presence of *ssc-miR-101* in sows'

endometrium might be predicting defective placentation and as a consequence, lower prolificacy levels.

After implantation, conceptus elongation becomes critical to establish an appropriate placental surface area that ensures successful embryo and foetal survival along gestation. This elongation is characterized by a morphological rearrangement of the conceptus trophectoderm from spherical to tubular shape. The porcine parathyroid hormone-like hormone (*PTHLH*) gene has been associated to this structural reorganization. Previous studies performed in our Iberian x Meishan F_2 intercross demonstrated that this gene maps into a significant QTL for teat number in SSC5. It was proposed as a candidate gene for this trait and it was demonstrated that it is involved in nipple formation during embryogenesis and nipple development during pregnancy and lactation. Our luciferase reporter assay indicates that *ssc-miR-144-3p* is able to down-regulate *PTHLH* expression in a relevant 69%. Again, previous transcriptomic evidences revealed significantly increased levels of this gene in the endometrium of high prolificacy samples, which is in agreement with the mentioned functions of this gene in successful embryo survival. In this particular case, RNA-seq results also revealed an increment of *ssc-miR-144-3p* in the endometrium of high prolificacy samples, which at first, could seem contradictory. However, despite miRNAs are known to repress gene expression, some studies recently revealed that they can also activate gene expression direct or indirectly depending on the cell type, conditions and cofactors (Valinezhad Orang et al. 2014). This versatility in their regulatory function allows the cells to quickly adapt to the changing conditions that take place in each tissue, avoiding an unnecessary waste of energy to maintain their state. As luciferase reporter assay has been performed in a different cell type, we could be also observing this phenomenon. Another feasible explanation to our observations is that microRNAs can simultaneously target several transcripts through cooperative/combinatorial targeting (Nunez et al. 2013). Thus, although *ssc-miR-144-3p* can be specifically down-regulating *PTHLH* expression as we observe in the luciferase assay, decreased expression levels of this gene could be masked by the action of some other miRNAs present in sows' endometrium. Moreover, it is also possible that miRNA targeting have not yet took place by the time endometrial samples were collected to perform RNA-sequencing, as these non-coding regulators present a spatially and temporally organized accumulation (Zheng et al. 2011).

Finally, we have analyzed the interaction between Adrenomedullin (*ADM*) and *ssc-miR-181d-5p*. This progesterone-dependent gene plays a relevant role in pregnancy establishment. This gene is an hypoxia-induced vasodilator peptide with a high expression in reproductive tissues such as uterine endometrium (Hague et al. 2000), fetal membranes(Trollmann et al. 2002) and placenta (Minegishi et al. 1999). Several animal models and also humans have been used to study the relation of *ADM* to fertility and implantations. In humans, A*DM* levels increase approximately in the maternal plasma of normal pregnancies compared with early pregnancies (Lenhart & Caron 2012). In rats, this gene has been associated to embryonic lethality and several pregnancy complications (Fritz-Six et al. 2008), in the regulation of progesterone production by the corpus luteum and in the transport of the embryo to the uterus (Liao et al. 2011). Based on this results, it is clear that *ADM* represents a determinant gene to ensure successful pregnancies and that any alteration of its expression level might probably lead to low reproductive efficiencies. Although transcriptomic data shows a perfect negative correlation between the expression of this gene and its target miRNA, we could not observe a down-regulation of *ADM* upon *ssc-miR-181d-5p* mimics.

To our knowledge, these results represent one of the first evidences of the miRNA-mediated regulation of key genes involved in porcine reproduction, functionally validating the effect that miRNA:mRNA interaction could porcine reproductive efficiency.

REFERENCES

Agarwal, V., Bell, G. W., Nam, J.-W., & Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, *4*. doi:10.7554/eLife.05005

Altmäe, S., Martinez-Conejero, J. A., Esteban, F. J., Ruiz-Alonso, M., Stavreus-Evers, A., Horcajadas, J. A., & Salumets, A. (2013). MicroRNAs miR-30b, miR-30d, and miR-494 regulate human endometrial receptivity. *Reproductive Sciences (Thousand Oaks, Calif.)*, *20*(3), 308–17. doi:10.1177/1933719112453507

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14744438

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*(38), 15144–9. doi:10.1073/pnas.0705917104

Chen, X., Li, A., Chen, W., Wei, J., Fu, J., & Wang, A. (2014). Differential Gene Expression in Uterine Endometrium During Implantation in Pigs. *Biology of Reproduction*, *92*(2), 52. doi:10.1095/biolreprod.114.123075

Estella, C., Herrer, I., Moreno-Moya, J. M., Quiñonero, A., Martínez, S., Pellicer, A., & Simón, C. (2012). miRNA signature and Dicer requirement during human endometrial stromal decidualization in vitro. *PloS One*, *7*(7), e41080. doi:10.1371/journal.pone.0041080

Fernandez-Rodriguez, A., Munoz, M., Fernandez, A., Pena, R. N., Tomas, A., Noguera, J. L., … 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*(2), 299–307. doi:10.1095/biolreprod.110.085589

Fritz-Six, K. L., Dunworth, W. P., Li, M., & Caron, K. M. (2008). Adrenomedullin signaling is necessary for murine lymphatic vascular development. *The Journal of Clinical Investigation*, *118*(1), 40–50. doi:10.1172/JCI33302

Gould, S. J., & Subramani, S. (1988). Firefly luciferase as a tool in molecular and cell biology. *Analytical Biochemistry*, *175*(1), 5–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3072883

Hague, S., Zhang, L., Oehler, M. K., Manek, S., MacKenzie, I. Z., Bicknell, R., & Rees, M. C. (2000). Expression of the hypoxically regulated angiogenic factor adrenomedullin correlates with uterine leiomyoma vascular density. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *6*(7), 2808–14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10914728

Hassan, S. S., Romero, R., Pineles, B., Tarca, A. L., Montenegro, D., Erez, O., … Kim, C. J. (2010). MicroRNA expression profiling of the human uterine cervix after term labor and delivery. *American Journal of Obstetrics and Gynecology*, *202*(1), 80.e1–80.e8. doi:10.1016/j.ajog.2009.08.016

Hou, X., Tang, Z., Liu, H., Wang, N., Ju, H., & Li, K. (2012). Discovery of MicroRNAs associated with myogenesis by deep sequencing of serial developmental skeletal muscles in pigs. *PloS One*, *7*(12), e52123. doi:10.1371/journal.pone.0052123

Hu, S.-J., Ren, G., Liu, J.-L., Zhao, Z.-A., Yu, Y.-S., Su, R.-W., … Yang, Z.-M. (2008). MicroRNA expression and regulation in mouse uterus during embryo implantation. *The Journal of Biological Chemistry*, *283*(34), 23473–84. doi:10.1074/jbc.M800406200

Hughes, P. E., & Varley, M. A. (1980). *Reproduction in the pig.* (1st ed.). Massachusetts: Butterworth-Heinemann Ltd.

Jin, Y., Chen, Z., Liu, X., & Zhou, X. (2013). Evaluating the microRNA targeting sites by luciferase reporter gene assay. *Methods in Molecular Biology (Clifton, N.J.)*, *936*, 117–27. doi:10.1007/978-1-62703-083-0_10

Koscianska, E., Starega-Roslan, J., & Krzyzosiak, W. J. (2011). The role of Dicer protein partners in the processing of microRNA precursors. *PloS One*, *6*(12), e28548. doi:10.1371/journal.pone.0028548

Krawczynski, K., Najmula, J., Bauersachs, S., & Kaczmarek, M. M. (2014). MicroRNAome of Porcine Conceptuses and Trophoblasts: Expression Profile of microRNAs and Their Potential to Regulate Genes Crucial for Establishment of Pregnancy. *Biology of Reproduction*. doi:10.1095/biolreprod.114.123588

Lenhart, P. M., & Caron, K. M. (2012). Adrenomedullin and pregnancy: perspectives from animal models to humans. *Trends in Endocrinology & Metabolism*, *23*(10), 524–532. doi:10.1016/j.tem.2012.02.007

Liao, S. B., Ho, J. C., Tang, F., & O, W. S. (2011). Adrenomedullin increases ciliary beat frequency and decreases muscular contraction in the rat oviduct. *Reproduction (Cambridge, England)*, *141*(3), 367–72. doi:10.1530/REP-10-0230

Lin, F., Li, R., Pan, Z. X., Zhou, B., Yu, D. B., Wang, X. G., … Liu, H. L. (2012). miR-26b promotes granulosa cell apoptosis by targeting ATM during follicular atresia in porcine ovary. *PloS One*, *7*(6), e38640. doi:10.1371/journal.pone.0038640

Minegishi, T., Nakamura, M., Abe, K., Tano, M., Andoh, A., Yoshida, M., … Kangawa, K. (1999). Adrenomedullin and atrial natriuretic peptide concentrations in normal pregnancy and pre-eclampsia. *Molecular Human Reproduction*, *5*(8), 767–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10421805

Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., … Kato, I. (2006). The expression profile of microRNAs in mouse embryos. *Nucleic Acids Research*, *34*(6), 1765–71. doi:10.1093/nar/gkl096

Montenegro, D., Romero, R., Kim, S. S., Tarca, A. L., Draghici, S., Kusanovic, J. P., … 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*(1), 113–21. doi:10.1002/path.2463

Nunez, Y. O., Truitt, J. M., Gorini, G., Ponomareva, O. N., Blednov, Y. A., Harris, R. A., & Mayfield, R. D. (2013). Positively correlated miRNA-mRNA regulatory networks in mouse frontal cortex during early stages of alcohol dependence. *BMC Genomics*, *14*(1), 725. doi:10.1186/1471-2164-14-725

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*(48), 20828–33. doi:10.1073/pnas.1008301107

Revel, A., Achache, H., Stevens, J., Smith, Y., & Reich, R. (2011). MicroRNAs are associated with human embryo implantation defects. *Human Reproduction (Oxford, England)*, *26*(10), 2830–40. doi:10.1093/humrep/der255

Rothschild, M. F. (1996). Genetics and reproduction in the pig. *Animal Reproduction Science*, *42*(1-4), 143–151. doi:10.1016/0378-4320(96)01486-8

Sha, A.-G., Liu, J.-L., Jiang, X.-M., Ren, J.-Z., Ma, C.-H., Lei, W., … Yang, Z.-M. (2011). Genome-wide identification of micro-ribonucleic acids associated with human endometrial receptivity in natural and stimulated cycles by deep sequencing. *Fertility and Sterility*, *96*(1), 150–155.e5. doi:10.1016/j.fertnstert.2011.04.072

Siomi, H., & Siomi, M. C. (2010). Posttranscriptional regulation of microRNA biogenesis in animals. *Molecular Cell*, *38*(3), 323–32. doi:10.1016/j.molcel.2010.03.013

Spötter, A., & Distl, O. (2006). Genetic approaches to the improvement of fertility traits in the pig. *Veterinary Journal (London, England : 1997)*, *172*(2), 234–47. doi:10.1016/j.tvjl.2005.11.013

Su, L., Liu, R., Cheng, W., Zhu, M., Li, X., Zhao, S., & Yu, M. (2014). Expression patterns of microRNAs in porcine endometrium and their potential roles in embryo implantation and placentation. *PloS One*, *9*(2), e87867. doi:10.1371/journal.pone.0087867

Tang, F., Kaneda, M., O'Carroll, D., Hajkova, P., Barton, S. C., Sun, Y. A., … Surani, M. A. (2007). Maternal microRNAs are essential for mouse zygotic development. *Genes & Development*, *21*(6), 644–8. doi:10.1101/gad.418707

Trollmann, R., Schoof, E., Beinder, E., Wenzel, D., Rascher, W., & Dotsch, J. (2002). Adrenomedullin gene expression in human placental tIssue and leukocytes: a potential marker of severe tIssue hypoxia in neonates with birth asphyxia. *European Journal of Endocrinology / European Federation of Endocrine Societies*, *147*(5), 711–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12444904

Valinezhad Orang, A., Safaralizadeh, R., & Kazemzadeh-Bavili, M. (2014). Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *International Journal of Genomics*, *2014*, 970607. doi:10.1155/2014/970607

Williams, K. C., Renthal, N. E., Condon, J. C., Gerard, R. D., & Mendelson, C. R. (2012). MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7529–34. doi:10.1073/pnas.1200650109

Williams, K. C., Renthal, N. E., Gerard, R. D., & Mendelson, C. R. (2012). The microRNA (miR)-199a/214 cluster mediates opposing effects of progesterone and estrogen on uterine contractility during pregnancy and labor. *Molecular Endocrinology (Baltimore, Md.)*, *26*(11), 1857–67. doi:10.1210/me.2012-1199

Xia, H.-F., Jin, X.-H., Cao, Z.-F., Hu, Y., & Ma, X. (2014). MicroRNA expression and regulation in the uterus during embryo implantation in rat. *The FEBS Journal*, *281*(7), 1872–91. doi:10.1111/febs.12751

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*(10), 3941–51. doi:10.1210/en.2011-1147

Yuan, Y., Liu, B., Xie, P., Zhang, M. Q., Li, Y., Xie, Z., & Wang, X. (2015). Model-guided quantitative analysis of microRNA-mediated regulation on competing endogenous RNAs using a synthetic gene circuit. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(10), 3158–63. doi:10.1073/pnas.1413896112

Zhao, Y., Lin, J., Xu, B., Hu, S., Zhang, X., & Wu, L. (2014). MicroRNA-mediated repression of nonsense mRNAs. *eLife*, *3*, e03032. doi:10.7554/eLife.03032

Zheng, G., Cochella, L., Liu, J., Hobert, O., & Li, W. (2011). Temporal and Spatial Regulation of MicroRNA Activity with Photoactivatable Cantimirs. *ACS Chemical Biology*, *6*(12), 1332–1338. doi:10.1021/cb200290e

4. GENERAL DISCUSSION

4. GENERAL DISCUSSION

Improvement of reproductive traits, which represent a high economic impact and understanding the complexity of those mechanisms involved in successful reproduction has been challenging through several years. The low heritability of these traits and the fact of being regulated by a complex network of interacting genes has been a limiting factor to successfully select individuals. Development of recent disciplines such as genomics or transcriptomics has served as a powerful tool in the study of complex traits, approaching researchers and breeders to the biological bases of reproductive success.

Reproduction in mammals is a highly complex process in which many events take place synchronously. It is a process that can be very different depending on the species, but always comprises several physiological, molecular and structural changes. Understanding those changes involved in pregnancy establishment is essential to increase reproductive efficiencies; however, there is a complex network of interacting genes involved. In pigs, the main limitation for increasing litter size is prenatal mortality which represent around 20–30% (days 10–30 of gestation) and 10–15% (days 50–70 of gestation). Recent evidences have indicated that prenatal loss in pigs results mainly from the decreased placental efficiency and uterine capacity (Vallet & Freking 2007; Ford et al. 2001). Therefore, as the uterus seems to play a key role in embryo implantation and litter size, in the present thesis we have explored the whole endometrial transcriptome profile of Iberian x Meishan F_2 pregnant sows at day 32 of their fifth gestation, with the aim to identify key differences in gene expression associated to swine reproductive efficiency.

In a first study, we performed an RNA-seq of endometrial samples with extreme phenotypes for reproductive efficiency, identifying 141 differentially expressed genes between high and low prolificacy sows. Subsequent functional enrichment analysis suggested that most of these genes were directly involved in pathways, such as progesterone and estrogen biosynthesis, immune recognition, membrane permeability, angiogenesis, transport of nutrients and signaling for pregnancy recognition, which are all highly relevant for pregnancy and embryonic development in swine. We also wanted to explore the regulatory mechanisms that could mediate this differential expression. A growing body of evidence demonstrates that miRNAs

represent an important mechanism of gene expression regulation, and several miRNAs are known to have key functions in many relevant biological processes involved in embryo formation, early development, and implantation (Y. Huang et al. 2011). Considering that, we also analyzed the miRNA expression profile in both extreme phenotypic groups predicting a total of 10 differentially expressed mature miRNAs between high and low prolificacy samples. Involvement of these small RNAs in the regulation of reproductive-related genes has been demonstrated by some authors in humans and other mammals (Teague et al. 2010; Logan & Hawkins 2015; Chegini 2010).

The main advantages of RNA-seq are the broad scope of genes being interrogated, its compatibility with allele and transcript specific RNA quantification, and the possibility of novel transcripts discovery. However, detection and quantification sensitivity of RNA-seq is very much depending on the read depth. Thus, validating our transcriptome sequencing results was essential in order to determine whether the observed expression differences in our samples were real or not. The benchmark technology for the detection of RNA levels is RT-qPCR. Despite being the highest sensitive RNA quantification technique, qPCR cost increases based on the number of genes being evaluated. For these reason, we performed a second study in which we wanted to assess the reproducibility of our RNA-seq results. To do so, among the 141 genes found differentially expressed between H and L groups, we selected those displaying the most extreme differences (FC ≥ 1.5) having a positive mapping into known reproductive QTLs and known to play a role in any relevant pathway related with reproduction based on both, enrichment and ingenuity pathway analysis (IPA) results. After applying these criteria, a smaller subset of 22 genes were analyzed by qPCR in 36 F_2 extreme individuals (H, n=18; L, n=18) obtaining significant differences for 13 genes between H and L samples. Considering their function, the validated genes are involved in the most relevant steps of porcine reproduction: Prostaglandins biosynthesis and pregnancy establishment and uterine receptivity and implantation.

Prostaglandins biosynthesis and pregnancy establishment

Prostaglandins (PGs) produced by the uterus play an essential role in luteolysis as well as in establishment of pregnancy in pigs and many other species (Blitek et al. 2006) (Bazer &

Thatcher 1977, McCracken *et al*. 1999, Waclawik *et al*. 2009*a*). Inhibition of PGs synthesis results in pregnancy failure (Spencer et al. 2004). The main PGs produced in the porcine endometrium are *PGE2* and *PGF2α*, with luteoprotective and luteolytic functions respectively. As they exert opposing functions, a tight control over their synthesis and secretion is critical either for the initiation of luteolysis or the maintenance of pregnancy. A rate-limiting enzyme in the production of PGs is the validated differentially expressed gene Prostaglandin endoperoxide synthase (*PTGS*; also known as prostaglandin G/H synthase or cyclooxygenase *COX2*). It catalyzes the conversion of arachidonic acid to *PGH2*, which is a common substrate for various prostaglandins. The conserved role of *PTGS2* in various species, including humans, has been widely discussed over the years and its key function to ensure reproductive success has been demonstrated through several previous studies.(Dey et al. 2004)'(S. Zhang et al. 2013). Considering that the production of prostaglandins directly contributes to the successful establishment of pregnancy, and that uterine receptivity to implantation is progesterone-dependent, a lack in the expression of this gene will directly affect the appropriate conceptus attachment. Based on this knowledge, we hypothesize that the observed underexpression of this gene in our low-prolificacy samples could be contributing to deficiencies in progesterone synthesis leading to embryonic deaths.

Another progesterone-dependent gene with key role in pregnancy establishment is the validated Adrenomedullin (*ADM*). This gene is an hypoxia-induced vasodilator peptide highly expressed in reproductive tissues such as uterine endometrium (Hague et al. 2000), fetal membranes (Trollmann et al. 2002) and placenta (Minegishi et al. 1999). Involvement of this gene in fertility and implantation has been studied in several animal models. In humans, it was observed that A*DM* levels increase approximately 5-fold in the maternal plasma of normal pregnancies compared with early pregnancies, especially at the earliest stages (Lenhart & Caron 2012). Similar results were found in mice by Fritz-Six and collaborators, who demonstrated that homozygous deletion of *Adm* causes embryonic lethality, and associating an altered *ADM* expression with several pregnancy complications (Fritz-Six et al. 2008). In rat, Lei *et. al.* showed that ovarian *Adm* expression appears to be involved in the regulation of progesterone production by the corpus luteum and Liao *et al*. pointed to a role of this gene in the regulation of embryo transport to the uterus (Liao et al. 2011). Significant over expression

of this gene in our high prolificacy samples might indicate a better outcome in the pregnancy establishment and the embryo development. But pregnancy establishment is not only associated to prostaglandins. In pigs, and many other mammals, this stage is characterized by the upregulation of pro-inflammatory factors, including cytokines, growth factors, and lipid mediators. The conceptus produces these inflammatory mediators (interferon γ and interferon δ, interleukins IL1B and IL6, and PGs) and maternal endometrium responds to these embryonic signals by enhancing further progesterone-induced uterine receptivity.

Uterine receptivity and implantation

Thus, a successful embryonic implantation needs a synchronized embryo-maternal dialogue. Heparan sulfate proteoglycans from the syndecan (*Sdc*) family such as the *SDCBP-2* gene found differentially expressed between high and low prolificacy samles, take part as co-receptors to help chemokines to bind with their innate receptors. This binding seems to mediate maternal acceptance towards embryo implantation (McEwan et al. 2009). Baston-Büst *et al.* observed that *Sdc-1* knock-down in human endometrial cells led to dramatic changes regarding cytokine expression profiles upon decidualization and embryonic contact (Baston-Büst et al. 2010). It is possible then, that the significant increase of *SDCBP-2* levels that we observe in our H prolificacy samples might support a better embryonic attachment and implantation due to the regulation of chemokine secretion of endometrial cells. The accumulation of chemokines, and other pro-inflammatory factors, is triggered by the Amyloid-beta peptide. Formation of amyloid-beta is catalyzed by gamma-secretase activation protein (*PION*, or *GSAP*) which selectively increases its production.

Simultaneously to maternal recognition of pregnancy, many structural changes take place in the uterine environment. A rapid transformation of trophoblast from spherical to tubular is essential to ensure successful implantation (Lala & Chakraborty 2003; Cha et al. 2012). Species displaying invasive implantation require a cell-to-cell communication through connexin proteins. Although porcine implantation is superficial, some authors have reported that endometrial cell-to-cell interaction may also be necessary for limiting trophoblast invasiveness or to develop specific channels that allow this superficial implantation (Wu et al. 2013). At this stage, the validated gene *MMP8* plays a key role. Proteins such as matrix metalloproteinase

(*MMP*) are a family of enzymes that use zinc-dependent catalysis to break down the components of the extracellular matrix (ECM) (Mousa et al. 2012; Schäfers et al. 2010; Wang et al. 2004). We hypothesize that the observed significant overexpression of this gene in our high-prolificacy samples may indicate a more efficient tissue reorganization to support the growing foetus.

Another relevant structural gene found differentially expressed in our extreme F_2 population is the Forkhead transcription factor *FOXA2*. This gene belong to a subfamily of Forkhead transcription factors that has been found to play an important role in early development, organogenesis, metabolism and homeostasis (Friedman & Kaestner 2006). Low-prolificacy samples show a decreased expression of this gene compared to those with high prolificacy, supporting our idea that an underexpression of this gene could be leading to defects in early development, affecting stages such as gastrulation or, later on, in embryo morphogenesis.

After implantation has occurred, embryonic growth and differentiation depends on the transport of nutrients and waste through the early vasculature. Thus, because of its importance the first functional organ system to develop in the vertebrate embryo is the cardiovascular system. The validated tissue Kallikrein gene (*KLK1*) is a member of a serine proteases family involved in many integral processes of early embryonic development which activates a wide range of substrates and growth factors (Stone et al. 2009). However, the fundamental roles of this gene seems to be vessel formation, vascular repair and robust arterialization(Stone et al. 2009). In humans, *KLK1* expression increases in first-trimester placentas, suggesting that it may participate in the establishment and maintenance of placental blood flow through vasodilatation and trophoblast invasion (Valdés et al. 2001; Luo et al. 2014). We observe a highly significant decrease of *KLK1* expression in low prolificacy samples. These results suggest that defects on the expression level of these gene may underlie serious reproductive conditions, probably due to defects in the ability of trophectoderm cells to fully invade the maternal uterine wall and remodel blood vessels (Lala & Chakraborty 2003; Chaddha et al. 2004). Trophectoderm cells play key role especially at the beginning of the attachment reaction, as the first cell type to interact with the blastocyst trophectoderm is the uterine luminal epithelium. The uterine luminal epithelium has to be conducive to blastocyst implantation and growth to ensure

a successful attachment, and this function is realized by a member of the the Kruppel-like factors (KLFs) family, the *KLF5* gene, which has been validated as differentially expressed between high and low prolificacy phenotypes. In its absence, trophectoderm development is defective resulting in developmental arrest at the blastocyst stage(Sun et al. 2012). The fact that this gene is over expressed in our high-prolificacy samples strengthens our idea of the important effect it may have on prolificacy levels and litter size control. This zinc finger-containing transcription factor, is known to regulate other cellular processes, including differentiation, proliferation, and apoptosis.(Parisi & Russo 2011)

Many other genes found differentially expressed in this study such as *CES1*, *FXYD3*, *PTHLH* and *SCNN1G* are also closely related with critical stages in embryo development at implantation level or later on in the survival of the embryo itself. For example, the porcine parathyroid hormone-like hormone (*PTHLH*) gene, which maps in a significant QTL for teat number in SSC5, has been proposed as a candidate gene for this trait in the same Iberian x Meishan F_2 intercross population that we are analyzing. Martínez-Giner and colaborators demonstrated that *PTHLH* was involved in nipple formation during embryogenesis and nipple development during pregnancy and lactation (Martínez-Giner et al. 2011). Another example is the FXYD Domain Containing Ion Transport Regulator 3 gene (FXYD3), which is also located into a known porcine reproductive QTL and has recently been proposed to be a candidate gene affecting litter size by influencing embryonic implantation (Chen et al. 2014). Along these experimental validations, we successfully validated 13 out of 22 predicted DEGs. We have observed that low abundant genes are those usually not validated by RT-qPCR. Some authors have suggested that this could be due to primers specificity and/or alternative splicing. It is possible that primers pick only one spliced variant reducing its abundance and making it very different to what you detect in RNA-Seq.

Besides exploring the coding genes being expressed in the endometrium of pregnant sows displaying extreme prolificacy phenotypes, in this first study we have explored the regulatory role that some candidate miRNAs exert in the expression of key reproductive-related genes and the possible effect that this has on litter size control.

MiRNA-mediated gene expression regulation

We performed an experimental validation of the expression level of 4 miRNAs, known to play key roles in reproductive processes: *ssc-miR-92a*, *ssc-miR-101*, *ssc-miR-133a* and *ssc-miR-181d*. In concordance with RNA-seq predictions, RT-qPCR results revealed and overexpression of *ssc-miR-101*, *ssc-miR-133a* and *ssc-miR-181d* in low prolificacy samples while *ssc-miR-92a* was overexpressed in high prolificacy samples. *MiR-92*, belongs to the *miR-17~92* cluster, known to regulate relevant processes for embryogenesis and pregnancy such as cardiac development, endothelial cell proliferation and angiogenesis(Bonauer & Dimmeler 2009). Loss and gain of function experiments showed that *miR-92a* can inhibit angiogenesis both, *in vitro* and *in vivo* (Bonauer et al. 2009) and that deletion of *miR-92a* induces critical skeletal defects in the developing embryo (Penzkofer et al. 2014). Thus, it is not surprising that we observe an overexpression of this miRNA in our high prolificacy samples, because its positive effect will confer a better outcome regarding pregnancy and embryo development.

Despite being predicted as differentially expressed by RNA-seq, when we performed RT-qPCR validations differences in their expression levels were below a FC of 1.5, and so, non-significant. When performing miRNA differential expression studies related to reproductive processes (L. Su et al. 2010; Balcells Ortega 2012), several authors have observed these small differences before and have demonstrated that even very small changes in microRNA expression levels (FC 1.5 to 2.5) could have a direct impact on their target genes. We hypothesize that this could be happening in our case, because despite these similar miRNA expression levels observed between both phenotypes, a significant correlation was found between the expression levels of validated genes *PTHLH*, *MMP8*, *PTGS2* and *SCNN1G*, and both *ssc-miR-133a* and *ssc-miR-92a*. Therefore, the finding of this significant correlation leads us to think that the observed differences, despite being low, may be biologically significant. One of the main reasons could be an insufficient sequencing depth in our small RNA libraries. Although no definitive guidelines on required sequence coverage have been given, this could be the reason why we are facing a bad agreement between the expression levels detected using both approaches.

With these two studies, we obtained a general view of the whole endometrial transcriptome and identified several potential candidate genes associated with critical steps involved in embryonic survival during the sow's gestation. We also described one of the possible regulatory mechanisms that could be responsible of the observed differences in the expression level of these key genes. To deepen the role of these regulatory mechanisms, we performed a third study in which we wanted to determine the mechanisms controlling miRNAs biosynthesis, in order to understand how this may affect their expression levels and therefore, their function as post-transcriptional regulators.

Control of miRNA expression

To date, the role of these small non-coding RNAs in maternal-fetal interactions through the regulation of uterine gene expression at the pre-implantation stage has been demonstrated(Bidarimath et al. 2014). Moreover, considering the capability of miRNAs to regulate multiple targets within the same pathway (Calin & Croce 2006), any alteration on the expression level of these small regulators could be associated to an alteration in embryo implantation and other reproductive diseases(Pan et al. 2007; Enquobahrie et al. 2011). Polymorphisms in either the primary or precursor form of a miRNA (miR-SNPs) have been proposed as a mechanism affecting mature microRNA expression either positively or negatively(Han et al. 2013). The results from our third study have demonstrated that polymorphisms identified at the precursor sequences of *ssc-mir-27a* [A/G] (2:65582002), *ssc-mir-29b-2* [A/G] (9:148552571) and *ssc-mir-106a* [G/C] (X: 126200101) are significantly associated with prolificacy phenotype in terms of EBV in our population.

Members belonging to the miR-29 family (which include miR-29a, miR-29b-1, miR-29b-2 and miR-29c) have been proposed as potent immune gene modulators(Liston et al. 2012). As previously discussed, upregulation of pro-inflammatory factors is necessary for implantation. During this period, there is an enrichment of immune cells, such as natural killer (NK) cells, T cells, B cells and macrophages at the maternal endometrium (Engelhardt 2002). These immune cells located at the maternal–fetal interface interact with foetal trophoblast cells allowing the growing foetus to develop its immunity (Erlebacher 2013) playing an important role in reproductive failures(Kwak-Kim et al. 2014). Moreover, besides its role as a mediator of the immune response, *miR-29b* has been found to be involved in the inhibition of trophoblast

differentiation (Kumar et al. 2013), gene reprogramming during endometrial stromal cell (ESC) decidualization (Qian et al. 2009) and in pre-eclampsia (Li et al. 2013). Our results showed that the presence of the variant identified at the precursor sequence of the *ssc-mir-29b-2* was associated with higher EBVs, being significantly higher in heterozygous individuals compared with homozygous individuals for both, the mutant and the reference allele. Subsequently, RT-qPCR results showed that the presence of the variant in this miRNA sequence significantly increases mature miRNA expression. These results are in agreement with the fact that miR-29 family seems to confer a better outcome in terms of both, immunity resistance of the embryo, and successful implantation. Thus, our hypothesis is that this increase on *ssc-miR-29b* expression promotes the inflammatory response necessary to stablish a successful implantation and thus, increases prolificacy levels.

We also identified a significant association between the polymorphism at the *ssc-mir-27a* and the prolificacy phenotype. *MiR-27a* has been studied in porcine placentas on days 30 and 90 of gestation identifying that it is able to target many genes that are key in reproductive processes, such as cell growth, trophoblast differentiation, angiogenesis and formation and maintenance of adherent junctions(L. Su et al. 2010). To our knowledge, this is the first miRNA in which the effect of a polymorphism in porcine litter size has been studied. This variant identified by Lei *et al.* was significantly associated with litter size in Large White and Meishan pigs(Lei et al. 2011). The variant identified in the precursor sequence of this miRNA in our population, involves a significant decrease of the EBV in homozygous samples. RT-qPCR results, however, show a significant increase of the mature miRNA expression on those homozygous samples. Considering that this miRNA targets several relevant reproduction-related genes, we hypothesize that higher expression levels of the mature form of *miR-27a* would result in a strong downregulation of these targets and as a consequence a decrease on prolificacy levels.

Finally, our validations have also focused on *ssc-mir-106a*. *MiR-106a-363* family, has been found to exert an inhibition of trophoblast differentiation (Kumar et al. 2013). Exactly as in *ssc-mir-27a*, we have observed that the presence of the variant has a significant impact on prolificacy levels and sows with CC genotype present decreased EBVs, and thus lower prolificacy levels. RT-qPCR results show that CC genotype involves an increase on the expression level of the mature miRNA. Because of its predicted role in the inhibition of

trophoblast, we hypothesize that higher expression values of this mature miRNA could have a negative effect on embryo attachment. Defects in the ability of trophectoderm cells to fully invade the maternal uterine wall and remodel blood vessels has been found to lead to defective embryo implantation (Lala & Chakraborty 2003; Chaddha et al. 2004) and this may explain the observed decrease on the EBVs of homozygous for the variant. With this study, we have demonstrated that variations found at the precursor level may be influencing the biosynthesis of the mature miRNA in a positive way, leading to the observed increase on their expression levels. However, this may alter the regulation process that these miRNAs exert on their target genes whose expression level plays key role in mechanisms involved in pig litter size variation. Surprisingly, in all cases we have observed that being heterozygous is associated with the best scenario in terms of prolificacy levels. It is possible that we were facing a heterosis or hybrid vigor. Heterosis has been widely studied in plant miRNAs and successfully applied in agricultural crops breeding, especially in maize. Although in 2002, Cassady and collaborators studied the effects of heterosis and recombination on pig reproductive traits, to date there are no previous evidences of this phenomenon in animal miRNAs and its underlying mechanism still remains poorly understood. Recent genomic and epigenetic studies suggest that heterosis might be explained by allelic interactions between parental genomes, leading to altered programming of genes that promote relevant traits of the hybrids (Chen 2013). Our hypothesis converges towards the idea of a cumulative positive effect of these mutations on the expression level of these miRNAs, causing a differential expression of a variety of genes that allow heterozygous to take advantages from progenitors.

After determining the miRNA expression profile of sows' endometrium, confirming the correlation between their expression levels and those from their putative targets and exploring the effect that single nucleotide variants have on their biosynthesis, we designed a fourth study to validate miRNAs effect on three of the most relevant candidate genes in porcine reproduction: *ADM*, *HTRA3*, *PTHLH* and *VEGFA*. Moreover, we wanted to stablish to what extent these interaction causes a downregulation of these genes affecting sows' prolificacy phenotypes. To achieve this goal, a luciferase reporter assay was performed. We observed a downregulation of *VEGFA* expression that was around 17% when using 33 nM of *ssc-miR-195-5p* mimics and 50ng of the reporter plasmid. Expression was reduced a 42% when using

100nM of of *ssc-miR-195-5p* mimics and 100ng of the reporter plasmid. VEGFA has been proposed as a candidate gene for litter size in pigs because of its role in embryonic implantation (Chen et al. 2014; Cordoba et al. 2014). Furthermore, single nucleotide polymorphism identified in *VEGFA* was shown to be associated with this trait also in pigs. A positive interaction between *HTRA3* and its targets *ssc-miR-101-3p* was also confirmed. In this case, gene expression decreased a 34% upon 100nM of *ssc-miR-101-3p* mimics and 100ng of the reporter plasmid. Although the role of this gene in porcine endometrium remains unknown, in humans it has been reported to inhibit TGF-β signalling in the endometrium and has been identified as a potential diagnostic marker for early detection of preeclampsia because it negatively regulates trophoblast invasion during placentation (Nie et al. 2006). However, the strongest downregulation was observed for *PTHLH* gene. Upon 100nM of *ssc-miR-144-3p* mimics, its expression was reduced in a 69% when using 100ng of the reporter plasmid.

The role of this gene in successful embryo survival along gestation has already been discussed in this thesis. Despite the validated miRNA-mediated down regulation of this gene by luciferase reporter assay, our previous transcriptomic evidences revealed significantly increased levels of both, *PTHLH* gene and its target miRNA *ssc-miR-144-3p* in the endometrium of high prolificacy samples, which at first, seemed contradictory. Although miRNAs are known to repress gene expression, some studies recently revealed they ability to activate gene expression depending on the cell type and conditions (Valinezhad Orang et al. 2014). This flexibility in their regulatory function would allow the cells to rapidly adapt to the changing conditions within each tissue. As we had performed the luciferase reporter assay in a different cell type, it could be possible to observe this phenomenon. Another feasible explanation to our observations is that microRNAs can target several transcripts simultaneously through cooperative targeting (Nunez et al. 2013). Thus, although *ssc-miR-144-3p* can be specifically down-regulating *PTHLH* expression, decreased gene expression levels could be masked by the action of some other miRNAs present in sows' endometrium. Furthermore, as these non-coding regulators present a spatially and temporally accumulation (Zheng et al. 2011), it is also possible that miRNA targeting had not yet took place by the time endometrial samples were collected to perform RNA-sequencing. Finally, we could not observe a down-regulation for *ADM* upon *ssc-miR-181d-5p* mimic. We have already discussed the role of this

gene in porcine reproduction and we have also validated its differential expression between high and low prolificacy samples through both, RNA-seq and RT-qPCR.

Altogether, our results confirm the impact that *ssc-miR-101-3p*, *ssc-miR-144-3p* and *ssc-miR-195-5p* exert on their respective target genes. Considering the relevant role of these targets in reproduction, these miRNAs could be a useful biomarker in the estimation of sows prolificacy levels. But besides miRNAs, we have also established which interactions exist between our validated candidate genes and other known regulatory molecules. We have discovered that there are two cytokines particularly capable of acting on the expression of some of the validated genes: the *ILK-1β* and the *TNF*.

Upstream regulators of gene expression: Beyond miRNAs

Upstream regulators such as transcription factors (TFs), growth factors (GFs) and many other molecules play critical roles as master regulators of gene expression. Investigating their involvement in a particular gene network or pathway can provide better clues on the underlying regulatory mechanisms that do mediate the observed differences in the expression of key genes in a particular biological context. We have discovered that cytokines *ILK-1β* and *TNF* are common upstream regulators of some relevant genes, which expression has been found differentially present between high and low prolificacy sows.

In reproductive biology, the role of these cytokines has been implicated in ovulation, menstruation, and embryo implantation, and pathological processes such as preterm delivery, and endometriosis (Dominguez et al. 2003; Simón et al. 1998). The interleukin 1 is a pro-inflammatory cytokine with multiple functions in a range of tissues (Dunne & O'Neill 2003). All components of the IL-1 system have been examined in the human endometrium and have been implicated as an important mediator of embryo implantation (Rossi et al. 2005; Healy et al. 2014). Simón C. and collaborators, demonstrated in mice, that *IL-1* receptor antagonist given before implantation significantly reduces the number of implanted embryos, indicating a role for IL-1 in embryo implantation (Simón et al. 1998).

The TNF is a pro-inflammatory cytokine that plays an important role in modulating the acute phase reaction. It was first discovered in amnion and placenta (Paradowska et al.), but many studies have demonstrated the presence of this cytokine and its receptors in the diverse human

reproductive tissues (Szarka et al. 2010). The TNF has been implicated in ovulation, corpus luteum formation and luteolysis, and it has been related to many endometrial and gestational diseases such as amniotic infections, recurrent spontaneous abortions, preeclampsia, preterm labour or endometriosis (Khan et al. 2011; Haider & Knöfler 2009; Hecht et al. 2011). Although these cytokines may be acting on the expression of our validated candidate genes, we haven't seen them differentially expressed between H and L groups.

This thesis represents one of the first descriptions of the mechanisms that affect embryonic survival in the pig, providing the knowledge to enhance fertility and reproductive health in this species. We have provided insight into the role of several candidate genes in litter size control and validated differences in their expression levels that can be attributed to extreme prolificacy phenotypes in our population. Because of the usefulness of the pig as a biomedical model and the parallelism in the function of these genes in humans, this study also provides a powerful tool to understand which genes are key in the process of embryo survival in mammals. Moreover we have explored the regulatory mechanisms that could regulate the expression of several potential candidate genes associated with critical steps involved in sow's gestation and how structural changes in miRNA precursor sequences could have an impact on mechanisms that mediate embryonic survival in the pig, providing the knowledge to enhance fertility by using miR-SNPs as biomarkers.

5. CONCLUSIONS

5. CONCLUSIONS

1. Transcriptome comparison between the endometrium of pregnant sows with extreme prolificacy phenotypes at day 32 of their gestation revealed the existence of 141 differentially expressed genes and 10 differentially expressed mature miRNAs.

2. Pathway analysis of differentially expressed genes showed that the main pathways in which these genes participate were female pregnancy, maternal placenta development and decidualization, which represent key processes for successful embryo implantation and development.

3. Among the 141 genes predicted as differentially expressed by RNA-seq, 22 candidates known to be involved in reproduction, displaying FC > 1.5 and having a positive mapping into known reproductive QTLs were selected for RT-qPCR validation. Significant expression differences were validated for 12 of them (*ADM*, *CES1*, *FXYD3*, *IHH*, *KLF5*, *KLK1*, *MMP8*, *PION*, *PTGS2*, *PTHLH*, *SCNN1G* and *SDCBP2*).

4. Among the 10 mature miRNAs predicted as differentially expressed by RNA-seq, 4 candidates known to be involved in the regulation of reproductive-related genes were selected for RT-qPCR validation. Similar expression levels were observed for all four miRNAs (*ssc-miR-92a,-101,-133a* and *-181d*). However, there was a significant correlation between the expression level of ssc-miR-92a and ssc-miR-133a and the validated genes *MMP8*, *PTGS2*, *PTHLH* and *SCNN1G*.

5. As single nucleotide polymorphisms (SNPs) are one of the mechanisms that could alter the expression level of miRNAs, a functional characterization of 9 reproduction-related pre-miRNAs was performed identifying 13 variants. Variants in 3 of these miRNAs (*ssc-mir-27a*,*-29b-2* and *-106*) were found to be directly associated with sow's EBVs.

6. We confirmed that these three variants cause significant differences in the expression level of the mature miRNA, being significantly higher in homozygous sows for the variant allele (for *ssc-mir-27a, s*ows with the GG genotype, for *ssc-mir-29b-2,* sows with the GG genotype and for *ssc-mir-106a, sows* with the CC genotype).

7. Finally, we have functionally validated that miRNA:mRNA interactions constitute one of the major mechanisms regulating key genes involved in pig litter size variation, demonstrating in cell culture that the interaction of *ssc-miR-101*, *-144* and *195* with their respective target genes *HTRA3*, *PTHLH* and *VEGFA*, causes a down-regulation in their expression level of 34%, 69% and 17% respectively.

6. REFERENCES

Agarwal, V. et al., 2015. Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, 4. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4532895&tool=pmcentrez&rendertype=abstract [Accessed August 13, 2015].

Altmäe, S. et al., 2013. MicroRNAs miR-30b, miR-30d, and miR-494 regulate human endometrial receptivity. *Reproductive sciences (Thousand Oaks, Calif.)*, 20(3), pp.308–17. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4077381&tool=pmcentrez&rendertype=abstract [Accessed March 5, 2015].

Andersson, L., 2001. Genetic dissection of phenotypic diversity in farm animals. *Nature reviews. Genetics*, 2(2), pp.130–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11253052 [Accessed September 29, 2015].

Andersson, L. et al., 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science (New York, N.Y.)*, 263(5154), pp.1771–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8134840 [Accessed August 27, 2015].

Anon, 2010. *Genomics: Essential Methods*, John Wiley & Sons. Available at: https://books.google.com/books?id=eduQHGLj7hcC&pgis=1 [Accessed September 30, 2015].

Archibald, A.L. et al., 1995. The PiGMaP consortium linkage map of the pig (Sus scrofa). *Mammalian genome : official journal of the International Mammalian Genome Society*, 6(3), pp.157–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7749223 [Accessed September 29, 2015].

Bagga, S. et al., 2005. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, 122(4), pp.553–63. Available at: http://www.sciencedirect.com/science/article/pii/S0092867405008019 [Accessed August 14, 2015].

Bail, S. et al., 2010. Differential regulation of microRNA stability. *RNA*, 16(5), pp.1032–1039. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2856875&tool=pmcentrez&rendertype=abstract [Accessed October 2, 2015].

Balcells, I., Castelló, A., et al., 2011. Analysis of porcine MUC4 gene as a candidate gene for prolificacy QTL on SSC13 in an Iberian × Meishan F2 population. *BMC genetics*, 12(1), p.93. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3224777&tool=pmcentrez&rendertype=abstract [Accessed January 28, 2013].

Balcells, I. et al., 2011. Sequencing and gene expression of the porcine ITIH SSC13 cluster and its effect on litter size in an Iberian × Meishan F2 population. *Animal reproduction science*, 128(1-4), pp.85–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21992966 [Accessed January 28, 2013].

Balcells, I., Cirera, S. & Busk, P.K., 2011. Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC biotechnology*, 11, p.70. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3135530&tool=pmcentrez&rendertype=abstract [Accessed November 13, 2014].

Balcells Ortega, I., 2012. *Deciphering the genetic architecture of prolificacy related traits in an experimental Iberian x Meishan F2 intercross*. Universitat Autònoma de Barcelona. Available at: http://www.tdx.cat/handle/10803/96517 [Accessed July 9, 2015].

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2), pp.281–97. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14744438 [Accessed July 17, 2014].

Baston-Büst, D.M. et al., 2010. Syndecan-1 knock-down in decidualized human endometrial stromal cells leads to significant changes in cytokine and angiogenic factor expression patterns. *Reproductive biology and endocrinology : RB&E*, 8, p.133. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2988802&tool=pmcentrez&rendertype=abstract [Accessed July 2, 2015].

Bauersachs, S. & Wolf, E., 2012. Transcriptome analyses of bovine, porcine and equine endometrium during the pre-implantation phase. *Animal reproduction science*, 134(1-2), pp.84–94. Available at: http://www.sciencedirect.com/science/article/pii/S037843201200245X [Accessed October 26, 2014].

Bazer, F.W., 2013. Pregnancy recognition signaling mechanisms in ruminants and pigs. *Journal of animal science and biotechnology*, 4(1), p.23. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3710217&tool=pmcentrez&rendertype=abstract [Accessed May 29, 2014].

Bazer, F.W. et al., 2011. Uterine receptivity to implantation of blastocysts in mammals. *Frontiers in bioscience (Scholar edition)*, 3, pp.745–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21196409 [Accessed January 15, 2015].

Bell, J.T. & Spector, T.D., 2011. A twin approach to unraveling epigenetics. *Trends in genetics : TIG*, 27(3), pp.116–25. Available at: http://www.sciencedirect.com/science/article/pii/S0168952510002428 [Accessed September 3, 2015].

Bentley, D.R. et al., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), pp.53–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2581791&tool=pmcentrez&rendertype=abstract [Accessed July 9, 2014].

Bertoldo, M.J. et al., 2012. Seasonal variation in the ovarian function of sows. *Reproduction, fertility, and development*, 24(6), pp.822–34. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22781933 [Accessed September 29, 2015].

Bidanel, 2015. Biology and Genetics of reproduction. *The Genetics of the Pig*. Available at: http://prodinra.inra.fr/?locale=en#!ConsultNotice:42313 [Accessed September 29, 2015].

Bidanel, J., 1993. Estimation of crossbreeding parameters between Large White and Meishan porcine breeds. III. Dominance and epistatic components of heterosis on reproductive traits. *Genetics Selection Evolution*, 25(3), p.263. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2710353/ [Accessed December 15, 2014].

Bidarimath, M. et al., 2014. MicroRNAs, immune cells and pregnancy. *Cellular & molecular immunology*, 11(6), pp.538–47. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4220836&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Blitek, A. et al., 2006. Expression of cyclooxygenase-1 and -2 in the porcine endometrium during the oestrous cycle and early pregnancy. *Reproduction in domestic animals = Zuchthygiene*, 41(3), pp.251–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16689891 [Accessed January 15, 2015].

Bonauer, A. et al., 2009. MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science (New York, N.Y.)*, 324(5935), pp.1710–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19460962 [Accessed July 9, 2014].

Bonauer, A. & Dimmeler, S., 2009. The microRNA-17-92 cluster: still a miRacle? *Cell cycle (Georgetown, Tex.)*, 8(23), pp.3866–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19887902 [Accessed November 17, 2014].

Bradford, G.E., 1979. Genetic variation in prenatal survival and litter size. *Journal of animal science*, 49 Suppl 2, pp.66–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/400778 [Accessed July 1, 2015].

Brennecke, J. et al., 2005. Principles of microRNA-target recognition. *PLoS biology*, 3(3), p.e85. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1043860&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

Brümmer, A. & Hausser, J., 2014. MicroRNA binding sites in the coding region of mRNAs: Extending the repertoire of post-transcriptional gene regulation. *BioEssays*, 36(6), pp.617–626. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24737341 [Accessed October 2, 2015].

Buermans, H.P.J. & den Dunnen, J.T., 2014. Next generation sequencing technology: Advances and applications. *Biochimica et biophysica acta*, 1842(10), pp.1932–1941. Available at: http://www.sciencedirect.com/science/article/pii/S092544391400180X [Accessed July 10, 2014].

Buske, B., Sternstein, I. & Brockmann, G., 2006a. QTL and candidate genes for fecundity in sows. *Animal reproduction science*, 95(3-4), pp.167–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16460893 [Accessed May 5, 2015].

Buske, B., Sternstein, I. & Brockmann, G., 2006b. QTL and candidate genes for fecundity in sows. *Animal reproduction science*, 95(3-4), pp.167–83. Available at: http://www.sciencedirect.com/science/article/pii/S0378432006000042 [Accessed June 18, 2014].

Bustin, S., 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, 25(2), pp.169–193. Available at: http://jme.endocrinology-journals.org/content/25/2/169 [Accessed September 22, 2015].

Bustin, S.A., 2004. *A-Z of Quantitative PCR* Igor F Tsigelny, ed., La Jolla, CA: IUL Biotechnology Series.

Bustin, S.A. et al., 2005. Quantitative real-time RT-PCR--a perspective. *Journal of molecular endocrinology*, 34(3), pp.597–601. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15956331 [Accessed July 16, 2014].

Bustin, S.A. et al., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*, 55(4), pp.611–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19246619 [Accessed July 9, 2014].

Cai, Y. et al., 2009. A brief review on the mechanisms of miRNA regulation. *Genomics, proteomics & bioinformatics*, 7(4), pp.147–54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20172487 [Accessed September 27, 2015].

Calin, G.A. & Croce, C.M., 2006. MicroRNA-cancer connection: the beginning of a new tale. *Cancer research*, 66(15), pp.7390–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16885332 [Accessed January 22, 2015].

Campbell, E.M.G., Nonneman, D. & Rohrer, G.A., 2003. Fine mapping a quantitative trait locus affecting ovulation rate in swine on chromosome 8. *Journal of animal science*, 81(7), pp.1706–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12854806 [Accessed November 13, 2014].

Carvalho, B.S. & Rustici, G., 2013. The challenges of delivering bioinformatics training in the analysis of high-throughput data. *Briefings in bioinformatics*, 14(5), pp.538–47. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3771233&tool=pmcentrez&rendertype=abstract [Accessed September 30, 2015].

Casellas, J. et al., 2008. Empirical Bayes factor analyses of quantitative trait loci for gestation length in Iberian × Meishan F2 sows. *Animal : an international journal of animal bioscience*, 2(2), pp.177–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22445010 [Accessed October 2, 2015].

Cassady, J.P. et al., 2001. Identification of quantitative trait loci affecting reproduction in pigs. *Journal of animal science*, 79(3), pp.623–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11263822 [Accessed November 13, 2014].

Cha, J., Sun, X. & Dey, S.K., 2012. Mechanisms of implantation: strategies for successful pregnancy. *Nature medicine*, 18(12), pp.1754–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23223073 [Accessed December 5, 2014].

Chaddha, V. et al., 2004. Developmental biology of the placenta and the origins of placental insufficiency. *Seminars in fetal & neonatal medicine*, 9(5), pp.357–69. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15691771 [Accessed July 3, 2015].

Chakrabarty, A. et al., 2007. MicroRNA regulation of cyclooxygenase-2 during embryo implantation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38), pp.15144–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1986627&tool=pmcentrez&rendertype=abstract.

Chang, C.Y.-Y. et al., 2011. MUC4 gene polymorphisms associate with endometriosis development and endometriosis-related infertility. *BMC medicine*, 9, p.19. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3052195&tool=pmcentrez&rendertype=abstract [Accessed October 2, 2015].

Chegini, N., 2010. Uterine microRNA signature and consequence of their dysregulation in uterine disorders. *Animal reproduction / Colegio Brasileiro de Reproducao Animal*, 7(3), pp.117–128. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3275910&tool=pmcentrez&rendertype=abstract [Accessed February 25, 2015].

Chen, C. et al., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic acids research*, 33(20), p.e179. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1292995&tool=pmcentrez&rendertype=abstract [Accessed July 9, 2014].

Chen, X. et al., 2014. Differential Gene Expression in Uterine Endometrium During Implantation in Pigs. *Biology of reproduction*, 92(2), p.52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25519183 [Accessed December 29, 2014].

Chen, Z.J., 2013. Genomic and epigenetic insights into the molecular bases of heterosis. *Nature reviews. Genetics*, 14(7), pp.471–82. Available at: http://dx.doi.org/10.1038/nrg3503 [Accessed May 27, 2015].

Cheung, V.G. et al., 2001. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature*, 409(6822), pp.953–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11237021 [Accessed September 29, 2015].

Coppieters, W. et al., 1995. Polymorphic CAC/T repetitive sequences in the pig genome 1. *Animal genetics*, 26(5), pp.327–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7486250 [Accessed September 29, 2015].

Cordoba, S. et al., 2014. Endometrial Gene Expression Profile from Pregnant Sows with Extreme Phenotype for Reproductive Capacity. In *34th International Society of Animal Genetics Conference*. Xian, p. 1.

Córdoba, S. et al., 2015. Endometrial gene expression profile of pregnant sows with extreme phenotypes for reproductive efficiency. *Scientific Reports*, 5, p.14416. Available at: http://www.nature.com/srep/2015/151005/srep14416/full/srep14416.html [Accessed October 5, 2015].

Corominas, J. et al., 2013. Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC genomics*, 14, p.843. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3879068&tool=pmcentrez&rendertype=abstract [Accessed December 15, 2014].

Dahan, O., Gingold, H. & Pilpel, Y., 2011. Regulatory mechanisms and networks couple the different phases of gene expression. *Trends in Genetics*, 27(8), pp.316–322. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21763027 [Accessed September 30, 2015].

Davies, W. et al., 1994. Characterization of microsatellites from flow-sorted porcine chromosome 13. *Mammalian genome : official journal of the International Mammalian Genome Society*, 5(11), pp.707–11. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7873881 [Accessed September 29, 2015].

Dey, S.K. et al., 2004. Molecular cues to implantation. *Endocrine reviews*, 25(3), pp.341–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15180948 [Accessed November 4, 2014].

Van Dijk, E.L. et al., 2014. Ten years of next-generation sequencing technology. *Trends in Genetics*, 30(9), pp.418–26. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25108476 [Accessed August 7, 2014].

Distl, O., 2007. Mechanisms of regulation of litter size in pigs on the genome level. *Reproduction in domestic animals = Zuchthygiene*, 42 Suppl 2, pp.10–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17688597 [Accessed November 17, 2014].

Dominguez, F., Pellicer, A. & Simon, C., 2003. The chemokine connection: hormonal and embryonic regulation at the human maternal-embryonic interface--a review. *Placenta*, 24 Suppl B, pp.S48–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14559030 [Accessed June 16, 2015].

Du, Z.-Q. et al., 2014. Identification of species-specific novel transcripts in pig reproductive tissues using RNA-seq. *Animal genetics*, 45(2), pp.198–204. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24450499 [Accessed June 30, 2015].

Dunne, A. & O'Neill, L.A.J., 2003. The interleukin-1 receptor/Toll-like receptor superfamily: signal transduction during inflammation and host defense. *Science's STKE : signal transduction knowledge environment*, 2003(171), p.re3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12606705 [Accessed June 16, 2015].

Echard, G. et al., 1992. The gene map of the pig (Sus scrofa domestica L.): a review. *Cytogenetics and cell genetics*, 61(2), pp.146–51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1395726 [Accessed September 29, 2015].

Ekimler, S. & Sahin, K., 2014. Computational Methods for MicroRNA Target Prediction. *Genes*, 5(3), pp.671–683. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4198924&tool=pmcentrez&rendertype=abstract [Accessed October 2, 2015].

Engelhardt, H., 2002. Conceptus Influences the Distribution of Uterine Leukocytes During Early Porcine Pregnancy. *Biology of Reproduction*, 66(6), pp.1875–1880. Available at: http://www.biolreprod.org/content/66/6/1875 [Accessed April 20, 2015].

Enquobahrie, D.A. et al., 2011. Placental microRNA expression in pregnancies complicated by preeclampsia. *American journal of obstetrics and gynecology*, 204(2), pp.178.e12–21. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3040986&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Erlebacher, A., 2013. Immunology of the maternal-fetal interface. *Annual review of immunology*, 31, pp.387–411. Available at: http://www.annualreviews.org/doi/full/10.1146/annurev-immunol-032712-100003 [Accessed February 7, 2015].

Ernst, C.W. & Steibel, J.P., 2013. Molecular advances in QTL discovery and application in pig breeding. *Trends in genetics : TIG*, 29(4), pp.215–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23498076 [Accessed August 27, 2015].

Estella, C. et al., 2012. miRNA signature and Dicer requirement during human endometrial stromal decidualization in vitro. *PloS one*, 7(7), p.e41080. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3401238&tool=pmcentrez&rendertype=abstract [Accessed February 25, 2015].

Esteve-Codina, A. et al., 2011. Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC genomics*, 12(1), p.552. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3221674&tool=pmcentrez&rendertype=abstract [Accessed November 9, 2012].

FAO, 2014. FAO's Animal Production and Health Division: Meat & Meat Products. *FAO's Animal Production and Health Division: Meat & Meat Products*. Available at: http://www.fao.org/ag/againfo/themes/en/meat/backgr_sources.html [Accessed September 29, 2015].

Fernandez-Rodriguez, A. et al., 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(2), pp.299–307. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20926806 [Accessed January 28, 2013].

Fernández-Rodríguez, A. et al., 2010. Analysis of candidate genes underlying two epistatic quantitative trait loci on SSC12 affecting litter size in pig. *Animal genetics*, 41(1), pp.73–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19793269 [Accessed October 2, 2015].

Ford, S.P., Vonnahme, K.A. & Wilson, M.E., 2001. Uterine capacity in the pig reflects a combination of uterine environment and conceptus genotype effects. *Journal of Animal Science*, 80(E-Suppl_1), pp.E66–E73. Available at: https://dl.sciencesocieties.org/publications/jas/abstracts/80/E-Suppl_1/JAN0080ES10E66 [Accessed July 9, 2015].

Franczak, A., Wojciechowicz, B. & Kotwica, G., 2013. Transcriptomic analysis of the porcine endometrium during early pregnancy and the estrous cycle. *Reproductive biology*, 13(3), pp.229–37. Available at: http://www.sciencedirect.com/science/article/pii/S1642431X13002490 [Accessed November 10, 2014].

Friedman, J.R. & Kaestner, K.H., 2006. The Foxa family of transcription factors in development and metabolism. *Cellular and molecular life sciences : CMLS*, 63(19-20), pp.2317–28. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16909212 [Accessed November 17, 2014].

Friedman, R.C. et al., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research*, 19(1), pp.92–105. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2612969&tool=pmcentrez&rendertype=abstract [Accessed July 11, 2014].

Fritz-Six, K.L. et al., 2008. Adrenomedullin signaling is necessary for murine lymphatic vascular development. *The Journal of clinical investigation*, 118(1), pp.40–50. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2147672&tool=pmcentrez&rendertype=abstract [Accessed May 25, 2015].

Furnes, B. & Schimenti, J., 2007. Fast forward to new genes in mammalian reproduction. *The Journal of physiology*, 578(Pt 1), pp.25–32. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2075132&tool=pmcentrez&rendertype=abstract [Accessed August 5, 2015].

Garcia, D.M. et al., 2011. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nature structural & molecular biology*, 18(10), pp.1139–46. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3190056&tool=pmcentrez&rendertype=abstract [Accessed September 4, 2014].

Goddard, M.E. & Hayes, B.J., 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature reviews. Genetics*, 10(6), pp.381–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19448663 [Accessed July 17, 2015].

Gould, S.J. & Subramani, S., 1988. Firefly luciferase as a tool in molecular and cell biology. *Analytical biochemistry*, 175(1), pp.5–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3072883 [Accessed August 24, 2015].

Govindarajan, B. & Gipson, I.K., 2010. Membrane-tethered mucins have multiple functions on the ocular surface. *Experimental eye research*, 90(6), pp.655–63. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2893012&tool=pmcentrez&rendertype=abstract [Accessed October 2, 2015].

Grabherr, M.G. et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), pp.644–52. Available at: http://dx.doi.org/10.1038/nbt.1883 [Accessed July 9, 2014].

Grimson, A. et al., 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular cell*, 27(1), pp.91–105. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3800283&tool=pmcentrez&rendertype=abstract [Accessed July 17, 2014].

Groenen, M.A. et al., 1995. Development and mapping of ten porcine microsatellite markers. *Animal genetics*, 26(2), pp.115–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7733492 [Accessed September 29, 2015].

Groenen, M.A.M. et al., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 491(7424), pp.393–8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3566564&tool=pmcentrez&rendertype=abstract [Accessed September 15, 2015].

Gu, T. et al., 2014. Endometrial gene expression profiling in pregnant Meishan and Yorkshire pigs on day 12 of gestation. *BMC genomics*, 15(1), p.156. Available at: http://www.biomedcentral.com/1471-2164/15/156 [Accessed November 13, 2014].

Hague, S. et al., 2000. Expression of the hypoxically regulated angiogenic factor adrenomedullin correlates with uterine leiomyoma vascular density. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 6(7), pp.2808–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10914728 [Accessed July 1, 2015].

Haider, S. & Knöfler, M., 2009. Human tumour necrosis factor: physiological and pathological roles in placenta and endometrium. *Placenta*, 30(2), pp.111–23. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2974215&tool=pmcentrez&rendertype=abstract [Accessed June 15, 2015].

Han, S.-J. et al., 2013. Kaposi's sarcoma-associated herpesvirus microRNA single-nucleotide polymorphisms identified in clinical samples can affect microRNA processing, level of expression, and silencing activity. *Journal of virology*, 87(22), pp.12237–48. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3807933&tool=pmcentrez&rendertype=abstract [Accessed April 20, 2015].

Harris, T.D. et al., 2008. Single-molecule DNA sequencing of a viral genome. *Science (New York, N.Y.)*, 320(5872), pp.106–9. Available at: http://www.sciencemag.org/content/320/5872/106.abstract [Accessed September 30, 2015].

Hassan, S.S. et al., 2010. MicroRNA expression profiling of the human uterine cervix after term labor and delivery. *American Journal of Obstetrics and Gynecology*, 202(1), pp.80.e1–80.e8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19889381 [Accessed March 26, 2015].

Healy, L.L., Cronin, J.G. & Sheldon, I.M., 2014. Endometrial cells sense and react to tissue damage during infection of the bovine endometrium via interleukin 1. *Scientific reports*, 4, p.7060. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4231323&tool=pmcentrez&rendertype=abstract [Accessed May 18, 2015].

Hecht, J.L. et al., 2011. Relationship Between Neonatal Blood Protein Concentrations and Placenta Histologic Characteristics in Extremely Low GA Newborns. *Pediatric research*, 69(1), pp.68–73. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3066075&tool=pmcentrez&rendertype=abstract.

Hill, W.G., 2014. Applications of population genetics to animal breeding, from wright, fisher and lush to genomic prediction. *Genetics*, 196(1), pp.1–16. Available at: /pmc/articles/PMC3872177/?report=abstract [Accessed August 28, 2015].

Hou, X. et al., 2012. Discovery of MicroRNAs associated with myogenesis by deep sequencing of serial developmental skeletal muscles in pigs. *PloS one*, 7(12), p.e52123. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3528764&tool=pmcentrez&rendertype=abstract [Accessed July 24, 2015].

Hu, S.-J. et al., 2008. MicroRNA expression and regulation in mouse uterus during embryo implantation. *The Journal of biological chemistry*, 283(34), pp.23473–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18556655 [Accessed April 30, 2014].

Hu, Y. et al., Two common SNPs in pri-miR-125a alter the mature miRNA expression and associate with recurrent pregnancy loss in a Han-Chinese population. *RNA biology*, 8(5), pp.861–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21788734 [Accessed February 4, 2015].

Hu, Z. et al., 2008. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *The Journal of clinical investigation*, 118(7), pp.2600–8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2402113&tool=pmcentrez&rendertype=abstract [Accessed February 24, 2015].

Hu, Z.-L. et al., 2013. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic acids research*, 41(Database issue), pp.D871–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3531174&tool=pmcentrez&rendertype=abstract [Accessed November 17, 2014].

Huang, R. et al., 2011. An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. *PloS one*, 6(11), p.e27288. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3213133&tool=pmcentrez&rendertype=abstract [Accessed October 28, 2012].

Huang, Y. et al., 2010. A study of miRNAs targets prediction and experimental validation. *Protein & cell*, 1(11), pp.979–86. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21153515 [Accessed December 7, 2012].

Huang, Y. et al., 2011. Biological functions of microRNAs: a review. *Journal of physiology and biochemistry*, 67(1), pp.129–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20981514 [Accessed November 7, 2012].

Hughes, P.E. & Varley, M.A., 1980. *Reproduction in the pig.* 1st ed., Massachusetts: Butterworth-Heinemann Ltd.

Huppertz, B. & Peeters, L.L.H., 2005. Vascular biology in implantation and placentation. *Angiogenesis*, 8(2), pp.157–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16211358 [Accessed July 3, 2015].

IDESCAT, 2015. Idescat. Demografia i qualitat de vida. Explotacions agràries amb ramaderia. Per espècies. Resultats. Available at: http://www.idescat.cat/economia/inec?tc=3&id=dd04&lang=es [Accessed September 29, 2015].

Jang, J.S. et al., 2011. Quantitative miRNA expression analysis using fluidigm microfluidics dynamic arrays. *BMC genomics*, 12(1), p.144. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3062620&tool=pmcentrez&rendertype=abstract [Accessed November 6, 2012].

Jin, Y. et al., 2013. Evaluating the microRNA targeting sites by luciferase reporter gene assay. *Methods in molecular biology (Clifton, N.J.)*, 936, pp.117–27. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3646406&tool=pmcentrez&rendertype=abstract [Accessed August 23, 2015].

Johnson, R.K., Nielsen, M.K. & Casey, D.S., 1999. Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. *Journal of animal science*, 77(3), pp.541–57. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10229349 [Accessed June 30, 2015].

Juskowiak, B., 2010. Nucleic acid-based fluorescent probes and their analytical potential. *Analytical and Bioanalytical Chemistry*, 399(9), pp.3157–3176. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3044240&tool=pmcentrez&rendertype=abstract [Accessed September 30, 2015].

Kaczmarek, M.M. et al., 2010. Seminal plasma affects prostaglandin synthesis in the porcine oviduct. *Theriogenology*, 74(7), pp.1207–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20615530 [Accessed September 29, 2015].

Karger, B.L. & Guttman, A., 2009. DNA sequencing by CE. *ELECTROPHORESIS*, 30(S1), pp.S196–S202. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2782523&tool=pmcentrez&rendertype=abstract [Accessed September 28, 2015].

Kawahara, Y. et al., 2007. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO reports*, 8(8), pp.763–9. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1978079&tool=pmcentrez&rendertype=abstract [Accessed January 20, 2013].

Khan, D.A., Ansari, W.M. & Khan, F.A., 2011. Pro/Anti-Inflammatory Cytokines in the Pathogenesis of Premature Coronary Artery Disease. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research*, 31(7), pp.561–567. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21323571 [Accessed July 1, 2011].

King, A.H. et al., 2003. Mapping quantitative trait loci affecting female reproductive traits on porcine chromosome 8. *Biology of reproduction*, 68(6), pp.2172–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12606397 [Accessed November 13, 2014].

Kloosterman, W.P. & Plasterk, R.H.A., 2006. The Diverse Functions of MicroRNAs in Animal Development and Disease. *Developmental Cell*, 11(4), pp.441–450. Available at: http://www.sciencedirect.com/science/article/pii/S1534580706004023 [Accessed October 2, 2015].

Koboldt, D.C. et al., 2013. The next-generation sequencing revolution and its impact on genomics. *Cell*, 155(1), pp.27–38. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3969849&tool=pmcentrez&rendertype=abstract [Accessed July 9, 2014].

De Koning, D.J. et al., 2001. Detection and characterization of quantitative trait loci for meat quality traits in pigs. *Journal of animal science*, 79(11), pp.2812–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11768109 [Accessed November 13, 2014].

Koscianska, E., Starega-Roslan, J. & Krzyzosiak, W.J., 2011. The role of Dicer protein partners in the processing of microRNA precursors. *PloS one*, 6(12), p.e28548. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3232248&tool=pmcentrez&rendertype=abstract [Accessed November 12, 2012].

Kotlabova, K., Doucha, J. & Hromadnikova, I., 2011. Placental-specific microRNA in maternal circulation--identification of appropriate pregnancy-associated microRNAs with diagnostic

potential. *Journal of reproductive immunology*, 89(2), pp.185–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21513988 [Accessed November 14, 2014].

Krawczynski, K. et al., 2014. MicroRNAome of Porcine Conceptuses and Trophoblasts: Expression Profile of microRNAs and Their Potential to Regulate Genes Crucial for Establishment of Pregnancy. *Biology of reproduction*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25472924 [Accessed January 15, 2015].

Krzymowski, T. & Stefańczyk-Krzymowska, S., 2008. The role of the endometrium in endocrine regulation of the animal oestrous cycle. *Reproduction in domestic animals = Zuchthygiene*, 43(1), pp.80–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18199263 [Accessed August 4, 2015].

Kumar, P. et al., 2013. The c-Myc-regulated microRNA-17~92 (miR-17~92) and miR-106a~363 clusters target hCYP19A1 and hGCM1 to inhibit human trophoblast differentiation. *Molecular and cellular biology*, 33(9), pp.1782–96. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3624183&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

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(2), pp.365–80. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3022285&tool=pmcentrez&rendertype=abstract [Accessed September 9, 2015].

Kwak-Kim, J. et al., 2014. Immunological modes of pregnancy loss: inflammation, immune effectors, and stress. *American journal of reproductive immunology (New York, N.Y. : 1989)*, 72(2), pp.129–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24661472 [Accessed April 11, 2015].

Lala, P.K. & Chakraborty, C., 2003. Factors regulating trophoblast migration and invasiveness: possible derangements contributing to pre-eclampsia and fetal injury. *Placenta*, 24(6), pp.575–87. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12828917 [Accessed July 3, 2015].

Larson, G. et al., 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 107(17), pp.7686–91. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2867865&tool=pmcentrez&rendertype=abstract [Accessed September 29, 2015].

Lee, Y., 2002. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, 21(17), pp.4663–4670. Available at: http://emboj.embopress.org/content/21/17/4663.abstract [Accessed January 26, 2015].

Lee, Y. et al., 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425(6956), pp.415–9. Available at: http://dx.doi.org/10.1038/nature01957 [Accessed February 17, 2015].

Lei, B. et al., 2011. A SNP in the miR-27a gene is associated with litter size in pigs. *Molecular biology reports*, 38(6), pp.3725–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21104015 [Accessed January 28, 2013].

Lenhart, P.M. & Caron, K.M., 2012. Adrenomedullin and pregnancy: perspectives from animal models to humans. *Trends in Endocrinology & Metabolism*, 23(10), pp.524–532. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3380178&tool=pmcentrez&rendertype=abstract [Accessed July 1, 2015].

Levine, M. & Tjian, R., 2003. Transcription regulation and animal diversity. *Nature*, 424(6945), pp.147–51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12853946 [Accessed July 31, 2015].

Li, C. et al., Molecular characterisation of porcine miR-155 and its regulatory roles in the TLR3/TLR4 pathways. *Developmental and comparative immunology*, 39(1-2), pp.110–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22301067 [Accessed February 20, 2015].

Li, L., O, W.-S. & Tang, F., 2011. Adrenomedullin in rat follicles and corpora lutea: expression, functions and interaction with endothelin-1. *Reproductive Biology and Endocrinology*, 9(1), p.111. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3175455&tool=pmcentrez&rendertype=abstract [Accessed July 1, 2015].

Li, P. et al., 2013. microRNA-29b contributes to pre-eclampsia through its effects on apoptosis, invasion and angiogenesis of trophoblast cells. *Clinical science (London, England : 1979)*, 124(1), pp.27–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22716646 [Accessed March 26, 2015].

Liao, S.B. et al., 2011. Adrenomedullin increases ciliary beat frequency and decreases muscular contraction in the rat oviduct. *Reproduction (Cambridge, England)*, 141(3), pp.367–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21173072 [Accessed July 1, 2015].

Lin, F. et al., 2012. miR-26b promotes granulosa cell apoptosis by targeting ATM during follicular atresia in porcine ovary. *PloS one*, 7(6), p.e38640. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3380909&tool=pmcentrez&rendertype=abstract [Accessed October 1, 2015].

Lin, S.-C.J. et al., 2010. Klf5 regulates lineage formation in the pre-implantation mouse embryo. *Development (Cambridge, England)*, 137(23), pp.3953–63. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2976279&tool=pmcentrez&rendertype=abstract [Accessed November 17, 2014].

Liston, A. et al., 2012. MicroRNA-29 in the adaptive immune system: setting the threshold. *Cellular and molecular life sciences : CMLS*, 69(21), pp.3533–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22971773 [Accessed April 20, 2015].

Liu, H.-C. et al., 2010. Current knowledge of microRNA characterization in agricultural animals. *Animal genetics*, 41(3), pp.225–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19968640 [Accessed July 14, 2015].

Liu, X. et al., 2014. Transcriptome Analysis of Porcine Thymus following Porcine Cytomegalovirus Infection. *PloS one*, 9(11), p.e113921. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4244220&tool=pmcentrez&rendertype=abstract [Accessed December 15, 2014].

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(4), pp.402–8. Available at: http://www.sciencedirect.com/science/article/pii/S1046202301912629 [Accessed July 9, 2014].

Logan, M. & Hawkins, S.M., 2015. Role of microRNAs in cancers of the female reproductive tract: insights from recent clinical and experimental discovery studies. *Clinical science (London, England : 1979)*, 128(3), pp.153–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25294164 [Accessed February 17, 2015].

López-Maury, L., Marguerat, S. & Bähler, J., 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature reviews. Genetics*, 9(8), pp.583–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18591982 [Accessed April 19, 2015].

Luo, Y. et al., 2014. Estrogen-related receptor γ serves a role in blood pressure homeostasis during pregnancy. *Molecular endocrinology (Baltimore, Md.)*, 28(6), pp.965–75. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4042076&tool=pmcentrez&rendertype=abstract [Accessed July 3, 2015].

MAGRAMA, 2014. Porcino - Sectores ganaderos - Producción y mercados ganaderos - Ganadería. *www.magrama.es*. Available at: http://www.magrama.gob.es/es/ganaderia/temas/produccion-y-mercados-ganaderos/sectores-ganaderos/porcino/default.aspx [Accessed September 29, 2015].

Margulies, M. et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), pp.376–80. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1464427&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

Martínez-Giner, M. et al., 2011. Expression study on the porcine PTHLH gene and its relationship with sow teat number. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 128(5), pp.344–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21906180 [Accessed July 11, 2014].

Mata, J., Marguerat, S. & Bähler, J., 2005. Post-transcriptional control of gene expression: a genome-wide perspective. *Trends in biochemical sciences*, 30(9), pp.506–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16054366 [Accessed July 29, 2015].

Mathew, D.J. et al., 2011. Uterine progesterone receptor expression, conceptus development, and ovarian function in pigs treated with RU 486 during early pregnancy. *Biology of reproduction*, 84(1), pp.130–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20864645 [Accessed November 14, 2014].

McEwan, M. et al., 2009. Cytokine regulation during the formation of the fetal-maternal interface: focus on cell-cell adhesion and remodelling of the extra-cellular matrix. *Cytokine & growth factor reviews*, 20(3), pp.241–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19487153 [Accessed July 2, 2015].

McKernan, K.J. et al., 2009. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Research*, 19(9), pp.1527–1541. Available at: http://genome.cshlp.org/content/19/9/1527 [Accessed September 30, 2015].

McLaren, W. et al., 2010. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics (Oxford, England)*, 26(16), pp.2069–70. Available at: http://bioinformatics.oxfordjournals.org/content/26/16/2069.abstract [Accessed July 16, 2014].

Mei, Q. et al., 2012. A facile and specific assay for quantifying microRNA by an optimized RT-qPCR approach. *PloS one*, 7(10), p.e46890. Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0046890 [Accessed September 30, 2015].

Mestdagh, P. et al., 2008. High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. *Nucleic acids research*, 36(21), p.e143. Available at: http://nar.oxfordjournals.org/content/36/21/e143.full [Accessed July 27, 2015].

Mewes, H.W. et al., 1997. Overview of the yeast genome. *Nature*, 387(6632 Suppl), pp.7–65. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9169865 [Accessed September 29, 2015].

Mihály, Z. & Gyorffy, B., 2011. [Next generation sequencing technologies (NGST) -- development and applications]. *Orvosi hetilap*, 152(2), pp.55–62. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21177232 [Accessed September 30, 2015].

Minegishi, T. et al., 1999. Adrenomedullin and atrial natriuretic peptide concentrations in normal pregnancy and pre-eclampsia. *Molecular human reproduction*, 5(8), pp.767–70. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10421805 [Accessed July 1, 2015].

Mineno, J. et al., 2006. The expression profile of microRNAs in mouse embryos. *Nucleic acids research*, 34(6), pp.1765–71. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1421506&tool=pmcentrez&rendertype=abstract [Accessed January 22, 2015].

Mishra, P.J. et al., 2008. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics. *Cell cycle (Georgetown, Tex.)*, 7(7), pp.853–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18414050 [Accessed February 17, 2015].

Montenegro, D. et al., 2009. Expression patterns of microRNAs in the chorioamniotic membranes: a role for microRNAs in human pregnancy and parturition. *The Journal of pathology*, 217(1), pp.113–21. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4160233&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Morin, R. et al., 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *BioTechniques*, 45(1), pp.81–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18611170 [Accessed September 30, 2015].

Mortazavi, A. et al., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), pp.621–8. Available at: http://dx.doi.org/10.1038/nmeth.1226 [Accessed July 11, 2014].

Mousa, A. a et al., 2012. Preeclampsia is associated with alterations in DNA methylation of genes involved in collagen metabolism. *The American journal of pathology*, 181(4), pp.1455–63. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3463634&tool=pmcentrez&rendertype=abstract [Accessed May 18, 2014].

Mullis, K. et al., 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor symposia on quantitative biology*, 51 Pt 1, pp.263–73. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3472723 [Accessed March 27, 2015].

Mutz, K.-O. et al., 2013. Transcriptome analysis using next-generation sequencing. *Current opinion in biotechnology*, 24(1), pp.22–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23020966 [Accessed July 11, 2014].

Nie, G. et al., 2006. Serine peptidase HTRA3 is closely associated with human placental development and is elevated in pregnancy serum. *Biology of reproduction*, 74(2), pp.366–74. Available at: http://www.biolreprod.org/content/74/2/366.abstract [Accessed September 17, 2015].

Noguera, J.L. et al., 2009. A bi-dimensional genome scan for prolificacy traits in pigs shows the existence of multiple epistatic QTL. *BMC genomics*, 10, p.636. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2812473&tool=pmcentrez&rendertype=abstract [Accessed January 28, 2013].

Nookaew, I. et al., 2012. A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae. *Nucleic acids research*, 40(20), pp.10084–97. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3488244&tool=pmcentrez&rendertype=abstract [Accessed July 11, 2014].

Nunez, Y.O. et al., 2013. Positively correlated miRNA-mRNA regulatory networks in mouse frontal cortex during early stages of alcohol dependence. *BMC genomics*, 14(1), p.725. Available at: http://www.biomedcentral.com/1471-2164/14/725 [Accessed October 1, 2015].

O'Leary, S. et al., 2004. Seminal plasma regulates endometrial cytokine expression, leukocyte recruitment and embryo development in the pig. *Reproduction (Cambridge, England)*, 128(2), pp.237–47. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15280563 [Accessed September 29, 2015].

Obad, S. et al., 2011. Silencing of microRNA families by seed-targeting tiny LNAs. *Nature genetics*, 43(4), pp.371–8. Available at: http://dx.doi.org/10.1038/ng.786 [Accessed October 2, 2015].

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, pp.87–102. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19848272 [Accessed October 2, 2015].

Ørom, U.A., Nielsen, F.C. & Lund, A.H., 2008. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular cell*, 30(4), pp.460–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18498749 [Accessed October 26, 2014].

Oshlack, A., Robinson, M.D. & Young, M.D., 2010. From RNA-seq reads to differential expression results. *Genome biology*, 11(12), p.220. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3046478&tool=pmcentrez&rendertype=abstract.

Østrup, E. et al., 2010. Differential endometrial gene expression in pregnant and nonpregnant sows. *Biology of reproduction*, 83(2), pp.277–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20393170 [Accessed September 29, 2015].

Pan, Q. et al., 2007. The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. *Molecular human reproduction*, 13(11), pp.797–806. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17766684 [Accessed October 30, 2012].

Paradowska, E., Blach-Olszewska, Z. & Gejdel, E., Constitutive and induced cytokine production by human placenta and amniotic membrane at term. *Placenta*, 18(5-6), pp.441–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9250707 [Accessed June 23, 2015].

Pareek, C.S., Smoczynski, R. & Tretyn, A., 2011. Sequencing technologies and genome sequencing. *Journal of applied genetics*, 52(4), pp.413–35. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3189340&tool=pmcentrez&rendertype=abstract [Accessed September 14, 2015].

Parisi, S. & Russo, T., 2011. Regulatory role of Klf5 in early mouse development and in embryonic stem cells. *Vitamins and hormones*, 87, pp.381–97. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22127252 [Accessed November 17, 2014].

Pennisi, E., 2012. Search for Pore-fection. *Science*, 336(6081), pp.534–537. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22556226 [Accessed September 30, 2015].

Penzkofer, D. et al., 2014. Phenotypic characterization of miR-92a-/- mice reveals an important function of miR-92a in skeletal development. A. Leri, ed. *PloS one*, 9(6), p.e101153. Available at: http://dx.plos.org/10.1371/journal.pone.0101153 [Accessed November 17, 2014].

Pfaffl, M.W., 2012. Quantification strategies in real-time polymerase chain reaction. In *Quantitative Real-time PCR in Applied Microbiology*. Norfolk, UK: Caister Academic Press, p. 53.

Pope, W.F. & First, N.L., 1985. Factors affecting the survival of pig embryos. *Theriogenology*, 23(1), pp.91–105. Available at: http://www.sciencedirect.com/science/article/pii/0093691X85900755 [Accessed June 15, 2015].

Pradervand, S. et al., 2009. Impact of normalization on miRNA microarray expression profiling. *RNA (New York, N.Y.)*, 15(3), pp.493–501. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2657010&tool=pmcentrez&rendertype=abstract [Accessed January 20, 2015].

Puig-Oliveras, A. et al., 2014. Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. *PloS one*, 9(6), p.e99720. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4057286&tool=pmcentrez&rendertype=abstract [Accessed December 15, 2014].

Qian, K. et al., 2009. Hsa-miR-222 is involved in differentiation of endometrial stromal cells in vitro. *Endocrinology*, 150(10), pp.4734–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19589872 [Accessed March 26, 2015].

Rajeevan, M.S. et al., 2001. Validation of Array-Based Gene Expression Profiles by Real-Time (Kinetic) RT-PCR. *The Journal of Molecular Diagnostics*, 3(1), pp.26–31. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1907344&tool=pmcentrez&rendertype=abstract [Accessed September 30, 2015].

Ramayo-Caldas, Y. et al., 2012. Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC genomics*, 13, p.547. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3478172&tool=pmcentrez&rendertype=abstract [Accessed October 31, 2012].

Ramirez, O. et al., 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 = Zuchthygiene*, 43(5), pp.542–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18312487 [Accessed October 2, 2015].

Rapaport, F. et al., 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome biology*, 14(9), p.R95. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4054597&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

Rathje, T.A., Rohrer, G.A. & Johnson, R.K., 1997. Evidence for quantitative trait loci affecting ovulation rate in pigs. *Journal of animal science*, 75(6), pp.1486–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9250508 [Accessed November 13, 2014].

Renthal, N.E. et al., 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(48), pp.20828–33. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2996411&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Revel, A. et al., 2011. MicroRNAs are associated with human embryo implantation defects. *Human reproduction (Oxford, England)*, 26(10), pp.2830–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21849299 [Accessed March 4, 2015].

Robinson, M.D. & Oshlack, A., 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology*, 11(3), p.R25. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2864565&tool=pmcentrez&rendertype=abstract.

Rodriguez, C. et al., 2005. QTL mapping for teat number in an Iberian-by-Meishan pig intercross. *Animal Genetics*, 36(6), p.050823030348002–??? Available at: http://www.ncbi.nlm.nih.gov/pubmed/16293122 [Accessed October 2, 2015].

Rohrer, G.A. et al., 1996. A comprehensive map of the porcine genome. *Genome research*, 6(5), pp.371–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8743988 [Accessed November 13, 2014].

Rohrer, G.A. et al., 1999. Identification of quantitative trait loci affecting female reproductive traits in a multigeneration Meishan-White composite swine population. *Journal of animal science*, 77(6), pp.1385–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10375216 [Accessed November 20, 2014].

Ross, J.W. et al., 2009. Identification of differential gene expression during porcine conceptus rapid trophoblastic elongation and attachment to uterine luminal epithelium. *Physiological genomics*, 36(3), pp.140–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19033546 [Accessed July 21, 2015].

Rossi, M. et al., 2005. Identification of genes regulated by interleukin-1beta in human endometrial stromal cells. *Reproduction (Cambridge, England)*, 130(5), pp.721–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16264101 [Accessed June 15, 2015].

Rothberg, J.M. et al., 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475(7356), pp.348–52. Available at: http://dx.doi.org/10.1038/nature10242 [Accessed July 9, 2014].

Rothschild, M.F., 1996. Genetics and reproduction in the pig. *Animal Reproduction Science*, 42(1-4), pp.143–151. Available at: http://www.sciencedirect.com/science/article/pii/0378432096014868 [Accessed June 18, 2014].

Rothschild, M.F., Hu, Z. & Jiang, Z., 2007. Advances in QTL mapping in pigs. *International journal of biological sciences*, 3(3), pp.192–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1802014&tool=pmcentrez&rendertype=abstract [Accessed November 13, 2014].

Rüegger, S. & Großhans, H., 2012. MicroRNA turnover: when, how, and why. *Trends in biochemical sciences*, 37(10), pp.436–46. Available at: http://www.sciencedirect.com/science/article/pii/S096800041200103X [Accessed September 10, 2015].

Ryan, B.M., Robles, A.I. & Harris, C.C., 2010. Genetic variation in microRNA networks: the implications for cancer research. *Nature reviews. Cancer*, 10(6), pp.389–402. Available at: http://dx.doi.org/10.1038/nrc2867 [Accessed November 25, 2014].

Rydhmer, L., 2000. Genetics of sow reproduction, including puberty, oestrus, pregnancy, farrowing and lactation. *Livestock Production Science*, 66(1), pp.1–12. Available at: http://www.sciencedirect.com/science/article/pii/S0301622699001700 [Accessed October 15, 2014].

Samborski, A. et al., 2013. Deep sequencing of the porcine endometrial transcriptome on day 14 of pregnancy. *Biology of reproduction*, 88(4), p.84. Available at: http://www.biolreprod.org/content/88/4/84.long [Accessed November 10, 2014].

Schäfers, M., Schober, O. & Hermann, S., 2010. Matrix-metalloproteinases as imaging targets for inflammatory activity in atherosclerotic plaques. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 51(5), pp.663–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20439506 [Accessed July 1, 2011].

Schmittgen, T.D. et al., 2000. Quantitative reverse transcription-polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. *Analytical biochemistry*, 285(2), pp.194–204. Available at: http://www.sciencedirect.com/science/article/pii/S0003269700947535 [Accessed July 29, 2015].

Schook, L.B. et al., 2005. Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome. *Comparative and functional genomics*, 6(4), pp.251–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2447480&tool=pmcentrez&rendertype=abstract [Accessed September 29, 2015].

Schwarzenbach, H. et al., 2015. Data Normalization Strategies for MicroRNA Quantification. *Clinical Chemistry*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26408530 [Accessed September 30, 2015].

Sha, A.-G. et al., 2011. Genome-wide identification of micro-ribonucleic acids associated with human endometrial receptivity in natural and stimulated cycles by deep sequencing. *Fertility and sterility*, 96(1), pp.150–155.e5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21601191 [Accessed March 26, 2015].

Shi, R. & Chiang, V.L., 2005. Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques*, 39(4), pp.519–25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16235564 [Accessed September 30, 2015].

Silió L., R.J.& T.M.A., 2001. La selección de cerdos ibéricos. In Buxadé C. and Daza A, ed. *Porcino ibérico: aspectos claves*. Madrid: S.A. MUNDI-PRENSA LIBROS, pp. 125–149.

Silió L; Rodriguez C; Rodrigáñez J; Toro MA, 2001. *PORCINO IBERICO ASPECTOS CLAVES*, Available at: http://www.libreriauniversitas.es/libros/PORCINO-IBERICO-ASPECTOS-CLAVES/LG060000646/978-84-7114-876-6 [Accessed December 15, 2014].

Simón, C. et al., 1998. Potential implications of chemokines in reproductive function: an attractive idea. *Journal of reproductive immunology*, 38(2), pp.169–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9730290 [Accessed June 16, 2015].

Siomi, H. & Siomi, M.C., 2010. Posttranscriptional regulation of microRNA biogenesis in animals. *Molecular cell*, 38(3), pp.323–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20471939 [Accessed January 19, 2015].

Siqueira, F.M. et al., 2014. Unravelling the transcriptome profile of the Swine respiratory tract mycoplasmas. *PloS one*, 9(10), p.e110327. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4198240&tool=pmcentrez&rendertype=abstract [Accessed December 15, 2014].

Smith, R.A. et al., A genetic variant located in miR-423 is associated with reduced breast cancer risk. *Cancer genomics & proteomics*, 9(3), pp.115–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22593246 [Accessed February 17, 2015].

Spencer, T.E. et al., 2004. Conceptus signals for establishment and maintenance of pregnancy. *Animal reproduction science*, 82-83, pp.537–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15271478 [Accessed January 6, 2015].

Spötter, A. & Distl, O., 2006a. Genetic approaches to the improvement of fertility traits in the pig. *Veterinary journal (London, England : 1997)*, 172(2), pp.234–47. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16426876 [Accessed November 13, 2014].

Spötter, A. & Distl, O., 2006b. Genetic approaches to the improvement of fertility traits in the pig. *Veterinary journal (London, England : 1997)*, 172(2), pp.234–47. Available at: http://www.sciencedirect.com/science/article/pii/S1090023305002959 [Accessed October 3, 2014].

Stone, O.A. et al., 2009. Critical role of tissue kallikrein in vessel formation and maturation: implications for therapeutic revascularization. *Arteriosclerosis, thrombosis, and vascular biology*, 29(5), pp.657–64. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2827862&tool=pmcentrez&rendertype=abstract [Accessed July 3, 2015].

Strachan, T. & Read, A.P., 2011. *Human molecular genetics* 4th ed., New York, NY: Garland Science/Taylor & Francis Group.

Streit, S. et al., 2009. Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. *Nature protocols*, 4(1), pp.37–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19131955 [Accessed August 24, 2015].

Su, L. et al., 2010. Differential expression of microRNAs in porcine placentas on days 30 and 90 of gestation. *Reproduction, fertility, and development*, 22(8), pp.1175–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20883642 [Accessed February 25, 2015].

Su, L. et al., 2014. Expression patterns of microRNAs in porcine endometrium and their potential roles in embryo implantation and placentation. *PloS one*, 9(2), p.e87867. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3914855&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Su, R.-W. et al., 2010. The integrative analysis of microRNA and mRNA expression in mouse uterus under delayed implantation and activation. *PloS one*, 5(11), p.e15513. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2993968&tool=pmcentrez&rendertype=abstract [Accessed October 1, 2014].

Sun, Q.-Y. & Nagai, T., 2003. Molecular mechanisms underlying pig oocyte maturation and fertilization. *The Journal of reproduction and development*, 49(5), pp.347–59. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14967910 [Accessed September 29, 2015].

Sun, X. et al., 2012. Kruppel-like factor 5 (KLF5) is critical for conferring uterine receptivity to implantation. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4), pp.1145–50. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3268277&tool=pmcentrez&rendertype=abstract [Accessed November 17, 2014].

Sun, X. et al., 2011. Microarray profiling for differential gene expression in PMSG-hCG stimulated preovulatory ovarian follicles of Chinese Taihu and Large White sows. *BMC genomics*, 12, p.111. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3047302&tool=pmcentrez&rendertype=abstract [Accessed July 21, 2015].

Szarka, A. et al., 2010. Circulating cytokines, chemokines and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex suspension array. *BMC immunology*, 11, p.59. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3014878&tool=pmcentrez&rendertype=abstract [Accessed June 17, 2015].

Tang, F. et al., 2007. Maternal microRNAs are essential for mouse zygotic development. *Genes & development*, 21(6), pp.644–8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1820938&tool=pmcentrez&rendertype=abstract [Accessed August 26, 2014].

Teague, E.M.C.O., Print, C.G. & Hull, M.L., 2010. The role of microRNAs in endometriosis and associated reproductive conditions. *Human reproduction update*, 16(2), pp.142–65. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19773286 [Accessed May 19, 2014].

Thorne, N., Inglese, J. & Auld, D.S., 2010. Illuminating Insights into Firefly Luciferase and Other Bioluminescent Reporters Used in Chemical Biology. *Chemistry & Biology*, 17(6), pp.646–657. Available at: http://www.sciencedirect.com/science/article/pii/S1074552110001973 [Accessed October 2, 2015].

Timoneda, O. et al., 2012. Determination of reference microRNAs for relative quantification in porcine tissues. *PloS one*, 7(9), p.e44413. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3438195&tool=pmcentrez&rendertype=abstract [Accessed November 17, 2014].

Tomas, 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(8), pp.1991–1998. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16864857 [Accessed October 2, 2015].

Tomás, A., Frigo, E., et al., 2006. An association study between polymorphisms of the porcine bone morphogenetic protein receptor type1beta(BMPR1B) and reproductive performance of Iberian x Meishan F2 sows. *Animal genetics*, 37(3), pp.297–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16734701 [Accessed October 2, 2015].

Tomás, A., Casellas, J., et al., 2006. Polymorphisms of the porcine dopamine beta-hydroxylase gene and their relation to reproduction and piglet survivability in an Iberian x Meishan F2 intercross. *Animal genetics*, 37(3), pp.279–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16734693 [Accessed October 2, 2015].

Trollmann, R. et al., 2002. Adrenomedullin gene expression in human placental tIssue and leukocytes: a potential marker of severe tIssue hypoxia in neonates with birth asphyxia. *European journal of endocrinology / European Federation of Endocrine Societies*, 147(5), pp.711–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12444904 [Accessed July 1, 2015].

Tse, C. & Capeau, J., [Real time PCR methodology for quantification of nucleic acids]. *Annales de biologie clinique*, 61(3), pp.279–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12805005 [Accessed September 30, 2015].

Tuggle, C.K., Wang, Y. & Couture, O., 2007. Advances in swine transcriptomics. *International journal of biological sciences*, 3(3), pp.132–52. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1802012&tool=pmcentrez&rendertype=abstract [Accessed August 27, 2015].

Valdés, G. et al., 2001. Tissue kallikrein in human placenta in early and late gestation. *Endocrine*, 14(2), pp.197–204. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11394637 [Accessed July 3, 2015].

Valencia-Sanchez, M.A. et al., 2006. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development*, 20(5), pp.515–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16510870 [Accessed July 14, 2014].

Valinezhad Orang, A., Safaralizadeh, R. & Kazemzadeh-Bavili, M., 2014. Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *International journal of genomics*, 2014, p.970607. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4142390&tool=pmcentrez&rendertype=abstract [Accessed October 1, 2015].

Vallet, J.L. & Freking, B.A., 2007. Differences in placental structure during gestation associated with large and small pig fetuses. *Journal of animal science*, 85(12), pp.3267–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17709791 [Accessed November 17, 2014].

Vanderhaeghe, C. et al., 2013. Non-infectious factors associated with stillbirth in pigs: a review. *Animal reproduction science*, 139(1-4), pp.76–88. Available at: http://www.sciencedirect.com/science/article/pii/S0378432013001048 [Accessed November 7, 2014].

Viganò, P. et al., 2003. Maternal-conceptus cross talk--a review. *Placenta*, 24 Suppl B, pp.S56–61. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14559031 [Accessed June 16, 2015].

Wang, H. et al., 2004. Functionally significant SNP MMP8 promoter haplotypes and preterm premature rupture of membranes (PPROM). *Human molecular genetics*, 13(21), pp.2659–69. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15367487 [Accessed May 19, 2014].

Wang, Z., Gerstein, M. & Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, 10(1), pp.57–63. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2949280&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

WELLS, L.J., 1946. Book review: Patterns of mammalian reproduction, by S. S. Asdell, Ithaca, 1946. *The Anatomical record*, 96(4), p.473. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20281501 [Accessed September 29, 2015].

Wessels, J.M. et al., 2013. The microRNAome of pregnancy: deciphering miRNA networks at the maternal-fetal interface. *PloS one*, 8(11), p.e72264. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3838410&tool=pmcentrez&rendertype=abstract [Accessed February 25, 2015].

Wightman, B., Ha, I. & Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, 75(5), pp.855–862. Available at: http://www.sciencedirect.com/science/article/pii/0092867493905304 [Accessed March 25, 2015].

Wilkie, P.J. et al., 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(6), pp.573–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10341088 [Accessed November 13, 2014].

Williams, K.C., Renthal, N.E., Condon, J.C., et al., 2012. MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), pp.7529–34. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3358858&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Williams, K.C., Renthal, N.E., Gerard, R.D., et al., 2012. The microRNA (miR)-199a/214 cluster mediates opposing effects of progesterone and estrogen on uterine contractility during pregnancy and labor. *Molecular endocrinology (Baltimore, Md.)*, 26(11), pp.1857–67. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3487626&tool=pmcentrez&rendertype=abstract [Accessed March 26, 2015].

Wu, D. et al., 2013. Molecular characterization and identification of the E2/P4 response element in the porcine HOXA10 gene. *Molecular and cellular biochemistry*, 374(1-2), pp.213–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23160802 [Accessed November 17, 2014].

Xia, H.-F. et al., 2014. MicroRNA expression and regulation in the uterus during embryo implantation in rat. *The FEBS journal*, 281(7), pp.1872–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24528955 [Accessed March 26, 2015].

Xu, L. & Tang, W., 2015. The associations of nucleotide polymorphisms in mir-196a2, mir-146a, mir-149 with lung cancer risk. *Cancer biomarkers : section A of Disease markers*, 15(1), pp.57–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25524943 [Accessed February 17, 2015].

Xu, S. et al., 2011. Micro-RNA378 (miR-378) regulates ovarian estradiol production by targeting aromatase. *Endocrinology*, 152(10), pp.3941–51. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3176644&tool=pmcentrez&rendertype=abstract [Accessed August 25, 2015].

Yang, J. et al., 2015. The Hedgehog signalling pathway in bone formation. *International journal of oral science*, 7(2), pp.73–9. Available at: http://dx.doi.org/10.1038/ijos.2015.14 [Accessed July 2, 2015].

Yang, R. et al., 2010. A genetic variant in the pre-miR-27a oncogene is associated with a reduced familial breast cancer risk. *Breast cancer research and treatment*, 121(3), pp.693–702. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19921425 [Accessed January 28, 2013].

Yang, W. et al., 2006. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature structural & molecular biology*, 13(1), pp.13–21. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2950615&tool=pmcentrez&rendertype=abstract [Accessed November 13, 2012].

Yerle, M. et al., 1995. The PiGMaP consortium cytogenetic map of the domestic pig (Sus scrofa domestica). *Mammalian genome : official journal of the International Mammalian Genome Society*, 6(3), pp.176–86. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7749224 [Accessed September 29, 2015].

Yu, Z. et al., 2007. Global analysis of microRNA target gene expression reveals that miRNA targets are lower expressed in mature mouse and Drosophila tissues than in the embryos. *Nucleic acids research*, 35(1), pp.152–64. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1802562&tool=pmcentrez&rendertype=abstract [Accessed November 13, 2014].

Yuan, Y. et al., 2015. Model-guided quantitative analysis of microRNA-mediated regulation on competing endogenous RNAs using a synthetic gene circuit. *Proceedings of the National Academy of Sciences of the United States of America*, 112(10), pp.3158–63. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4364186&tool=pmcentrez&rendertype=abstract [Accessed October 1, 2015].

Zhang, H. et al., 2013. Differential gene expression in the endometrium on gestation day 12 provides insight into sow prolificacy. *BMC genomics*, 14, p.45. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3610143&tool=pmcentrez&rendertype=abstract [Accessed November 14, 2014].

Zhang, S. et al., 2013. Deciphering the molecular basis of uterine receptivity. *Molecular reproduction and development*, 80(1), pp.8–21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23070972 [Accessed November 4, 2014].

Zhang, X. et al., 2015. Association of the miR-146a , miR-149 , miR-196a2 and miR-499 polymorphisms with susceptibility to pulmonary tuberculosis in the Chinese Uygur, Kazak and Southern Han populations. *BMC infectious diseases*, 15(1), p.41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25650003 [Accessed February 17, 2015].

Zhao, Y. et al., 2014. MicroRNA-mediated repression of nonsense mRNAs. *eLife*, 3, p.e03032. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4359369&tool=pmcentrez&rendertype=abstract [Accessed September 15, 2015].

Zheng, G. et al., 2011. Temporal and Spatial Regulation of MicroRNA Activity with Photoactivatable Cantimirs. *ACS Chemical Biology*, 6(12), pp.1332–1338. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3643504&tool=pmcentrez&rendertype=abstract [Accessed October 1, 2015].

Zhou, Q.-Y. et al., 2009. Detection of differentially expressed genes between Erhualian and Large White placentas on day 75 and 90 of gestation. *BMC genomics*, 10, p.337. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2724418&tool=pmcentrez&rendertype=abstract [Accessed November 14, 2014].

7. ANNEXES

7.1. Supplementary material from Study I: Analysis of gene expression differences between extreme prolificacy phenotypes

Table S1. Differential expressed genes found in mRNA libraries between both extreme phenotypes. All expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads – Mortazavi et al., 2008). Mean difference between both groups is represented as the log_2 transformed fold change (Log_2FC).

Gene name	Locus	High RPKM	Low RPKM	Log_2 FC	p-value	q-value[a]	Up regulation
Un-annotated	2:120052627-120052759	6159,840	1190,200	-2.372	0.0001	0.0129	High
ADM	2:52576214-52578540	165,926	35,818	-2.212	0.0004	0.0319	High
ANXA8	14:95804060-95825709	134,914	10,747	-3.650	0.0001	0.0078	High
ATP1B1	4:89529320-89550560	213,255	18,941	-3.493	0.0001	0.0078	High
BF	7:27771731-27777630	17,525	4,625	-1.922	0.0003	0.0282	High
CES1	6:27276810-27380644	88,817	10,330	-3.104	0.0001	0.0078	High
CST6	2:5395319-5396971	904,334	236,311	-1.936	0.0002	0.0173	High
CXCL16	12:54264905-54269592	185,155	36,204	-2.354	0.0002	0.0173	High
DF	2:77485871-77487847	79,457	21,607	-1.879	0.0002	0.0173	High
DPCD	14:122343533-122381270	81,013	18,670	-2.117	0.0006	0.0370	High
EGLN3	7:70667588-70695692	28,945	3,381	-3.098	0.0001	0.0078	High
ENSSSCG00000004703	1:142682190-142688725	162,847	29,120	-2.483	0.0001	0.0129	High
ENSSSCG00000010533	14:119116727-119128192	103,147	1,800	-5.840	0.0005	0.0360	High
ENSSSCG00000012427	X:69067109-69087469	1318,870	245,174	-2.427	0.0003	0.0282	High
ENSSSCG00000013976	2:55147384-55155430	175,303	5,964	-4.877	0.0001	0.0078	High
ENSSSCG00000026236	13:170261824-170278122	142,493	38,629	-1.883	0.0001	0.0129	High
ENSSSCG00000027404	1:3101176-3126386	59,013	12,192	-2.275	0.0008	0.0454	High
ENSSSCG00000027784	16:51499064-51507622	35,405	8,666	-2.030	0.0003	0.0262	High
ENSSSCG00000028525	2:44078430-44082946	135,308	17,341	-2.964	0.0004	0.0319	High
ENSSSCG00000028923	10:309238-337906	34,982	2,141	-4.030	0.0001	0.0078	High
EPHA1	18:7081287-7098452	22,152	6,946	-1.673	0.0003	0.0282	High
EYA2	17:54887015-55078442	30,370	7,571	-2.004	0.0002	0.0215	High
FXYD3	6:40111306-40118025	211,400	22,672	-3.221	0.0001	0.0078	High
GPR110	7:48466208-48513535	38,595	8,630	-2.161	0.0001	0.0078	High
HDC	1:135147966-135166937	20,811	1,531	-3.765	0.0003	0.0282	High
HMGCR	2:85967320-85990095	56,566	12,800	-2.144	0.0001	0.0129	High
HPGD	14:16908629-16921262	102,135	28,404	-1.846	0.0004	0.0319	High
HSD17B7	4:95571647-95594215	75,112	21,034	-1.836	0.0002	0.0173	High
KLF5	11:49867087-49870148	210,133	53,936	-1.962	0.0005	0.0360	High
LCN2	1:302600678-302605199	80,798	16,342	-2.306	0.0001	0.0078	High
MAPK4	1:110082309-110155111	44,643	7,117	-2.649	0.0001	0.0078	High
MMP8	9:37330737-37343663	68,489	8,639	-2.987	0.0001	0.0078	High
MSLN	3:41448661-41452015	40,196	4,376	-3.199	0.0002	0.0215	High
MST1R	13:35590300-35603228	28,392	7,735	-1.876	0.0002	0.0173	High

Gene name	Locus	High RPKM	Low RPKM	Log_2 FC	p-value	q-value[a]	Up regulation
MTMR11	4:108596807-108604443	38,519	11,268	-1.773	0.0002	0.0173	High
MUC1	4:103409754-103413841	148,217	22,219	-2.738	0.0003	0.0282	High
MUC4	13:143786442-143842402	44,827	10,970	-2.031	0.0004	0.0319	High
NOP56	17:37441072-37451309	219,218	36,853	-2.573	0.0005	0.0352	High
OVOL1	2:5630670-5642984	22,256	4,696	-2.245	0.0001	0.0078	High
PARP3	13:37272322-37280891	21,833	8,022	-1.445	0.0007	0.0433	High
PLA2G4A	9:140460879-140623439	17,008	4,535	-1.907	0.0003	0.0282	High
PSAT1	1:257809814-257835728	38,254	10,211	-1.906	0.0007	0.0414	High
PTGS2	9:140251515-140260362	109,913	6,574	-4.063	0.0001	0.0078	High
PTHLH	5:49160138-49172386	286,097	14,359	-4.316	0.0001	0.0078	High
RAB25	4:102703921-102713495	87,765	20,834	-2.075	0.0001	0.0078	High
SCNN1G	3:23444853-23480156	33,295	2,647	-3.653	0.0001	0.0078	High
SDCBP2	17:38419446-38437273	49,494	6,487	-2.932	0.0001	0.0078	High
SFN	6:77758723-77760027	66,169	11,685	-2.501	0.0001	0.0078	High
SGPP2	15:137861023-137905123	11,590	1,686	-2.781	0.0001	0.0078	High
SLC52A3	17:39142052-39160998	13,635	2,891	-2.238	0.0007	0.0433	High
SMOC1	7:99860647-99938608	42,954	6,162	-2.801	0.0001	0.0078	High
STAP2	2:74955494-74964488	24,900	4,142	-2.588	0.0003	0.0282	High
TMEM139	18:7211799-7214244	65,159	4,274	-3.930	0.0001	0.0078	High
TMEM79	4:102462253-102467485	28,115	5,337	-2.397	0.0006	0.0370	High
VEGF,VEGFA	7:44224281-44475316	992,547	60,231	-4.043	0.0001	0.0078	High
Un-annotated	2:76981417-76982104	34,069	112,967	1.729	0.0005	0.0360	Low
Un-annotated	3:10903133-10908516	12,831	46,010	1.842	0.0002	0.0173	Low
Un-annotated	3:19921984-19925331	20,020	83,411	2.059	0.0003	0.0262	Low
APOA1	9:49288614-49290784	18,310	88,047	2.266	0.0001	0.0078	Low
CEBPD	4:87368560-87610320	25,189	89,790	1.834	0.0008	0.0454	Low
CFL2	7:69919876-69924154	42,519	146,918	1.789	0.0005	0.0360	Low
CHRAC1	4:2425324-2427426	38,895	119,120	1.615	0.0004	0.0308	Low
CLEC3B	13:31097019-31106889	47,058	176,399	1.906	0.0002	0.0215	Low
CYP17A1	14:123773104-123779533	4,570	52,641	3.526	0.0001	0.0078	Low
DCLK2	8:83617784-83721850	14,485	51,711	1.836	0.0007	0.0433	Low
DPT	4:89969323-90001121	17,523	51,942	1.568	0.0005	0.0360	Low
ECHDC1	1:39547701-39593566	6,554	29,560	2.173	0.0001	0.0129	Low
ENDOD1	9:30923140-30952504	7,300	29,739	2.026	0.0001	0.0078	Low
ENPEP	8:119968857-120061196	6,394	49,590	2.955	0.0001	0.0078	Low
ENSSSCG00000000921	5:97776302-97793667	5,701	17,837	1.646	0.0006	0.0370	Low
ENSSSCG00000004572	1:121176825-121178352	2,404	46,442	4.272	0.0008	0.0454	Low
ENSSSCG00000004573	1:121348386-121353049	23,755	205,948	3.116	0.0001	0.0078	Low
ENSSSCG00000008627	3:133981946-134005017	12,734	51,289	2.010	0.0004	0.0319	Low
ENSSSCG00000010464	14:112652686-112657270	13,535	45,323	1.744	0.0001	0.0129	Low
ENSSSCG00000013152	2:12110927-12113060	7,494	235,166	4.972	0.0001	0.0129	Low
ENSSSCG00000017492	12:22869394-22876022	31,662	89,836	1.505	0.0009	0.0499	Low
ENSSSCG00000025083	1:108567669-108607804	37,358	232,497	2.638	0.0004	0.0319	Low
ENSSSCG00000026285	6:90674121-90703000	10,746	41,930	1.964	0.0004	0.0319	Low
ENSSSCG00000029421	1:268964059-268973843	7,241	41,506	2.519	0.0001	0.0129	Low

Gene name	Locus	High RPKM	Low RPKM	Log_2 FC	p-value	q-value[a]	Up regulation
EPC1	10:47431312-47541566	19,907	210,948	3.406	0.0001	0.0078	Low
FAM174B	7:92211760-92247931	7,952	26,642	1.744	0.0004	0.0308	Low
FOXA2	17:34053459-34056624	9,494	37,100	1.966	0.0001	0.0129	Low
GPER	3:828570-833135	3,927	24,583	2.646	0.0001	0.0078	Low
HBE1	9:5650627-5652665	20,916	1024,720	5.614	0.0001	0.0078	Low
HTRA3	8:4456175-4482055	19,965	77,677	1.960	0.0004	0.0319	Low
IHH	15:134122694-134129391	5,928	28,747	2.278	0.0002	0.0215	Low
JUNB	2:66505143-66507024	25,216	79,818	1.662	0.0004	0.0308	Low
KLF2	1:279056680-279061593	7,126	22,844	1.681	0.0007	0.0433	Low
MGP	5:61054165-61058328	157,765	502,327	1.671	0.0007	0.0414	Low
MME	13:103030829-103128447	26,332	85,476	1.699	0.0009	0.0499	Low
MMP23B	6:58231350-58233984	7,099	55,587	2.969	0.0001	0.0078	Low
MYEOV2	15:153946103-153950517	27,781	86,956	1.646	0.0007	0.0414	Low
NEXN	6:125500472-125693958	11,924	60,290	2.338	0.0001	0.0078	Low
PDK4	9:82625076-82638263	22,654	72,263	1.673	0.0005	0.0360	Low
PI15	4:66776707-66802989	9,080	30,478	1.747	0.0001	0.0078	Low
PION	9:113183480-113208758	2,916	14,502	2.314	0.0009	0.0499	Low
RGS5	4:94999825-95057119	23,315	131,159	2.492	0.0002	0.0173	Low
ROR2	14:3602842-3643321	6,403	18,823	1.556	0.0005	0.0360	Low
SAL1	1:284447109-284451960	4,702	37,998	3.015	0.0006	0.0391	Low
SH3BGR	13:213286139-213348206	14,455	58,620	2.020	0.0001	0.0078	Low
SLC24A4	7:120438866-120615704	7,877	34,496	2.131	0.0006	0.0370	Low
SST	13:134620965-134622407	12,591	155,456	3.626	0.0001	0.0129	Low
TM9SF2	11:75703401-75712182	36,818	144,519	1.973	0.0006	0.0370	Low
WFDC1	6:4754997-4789463	7,011	19,908	1.506	0.0006	0.0391	Low

Table S2. Differentially expressed genes found uniquely expressed in one of the prolificacy groups in mRNA libraries. All expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads). Shown *q-values* are Benjamini-Hochberg false discovery rate (FDR) corrected *p-values*. Mean difference between both groups is represented as the log_2 transformed fold change (Log_2FC).

Gene name	Locus	High RPKM	Low RPKM	*p-value*	*q-value*	Expression
un-annotated	1:305102825-305126687	0,000	22,438	0.0001	0.0078	low group
un-annotated	10:63437508-63437903	0,000	15,437	0.0003	0.0282	low group
un-annotated	11:86787469-86788850	0,000	21,723	0.0004	0.0308	low group
un-annotated	11:49276974-49322148	0,000	38,516	0.0005	0.0360	low group
un-annotated	11:306177-306599	0,000	11,421	0.0006	0.0370	low group
un-annotated	12:6776624-6776820	0,000	101,847	0.0005	0.0360	low group
un-annotated	12:13704609-13705014	0,000	24,158	0.0001	0.0078	low group
un-annotated	13:3656448-3695027	0,000	57,550	0.0001	0.0078	low group
un-annotated	13:34286207-34286740	0,000	30,630	0.0006	0.0391	low group
un-annotated	13:191013924-191014235	0,000	28,185	0.0002	0.0215	low group
un-annotated	15:117232153-117235744	0,000	17,595	0.0001	0.0078	low group
un-annotated	15:122380337-122380596	0,000	77,571	0.0001	0.0078	low group
un-annotated	16:77861325-77861582	0,000	36,153	0.0008	0.0480	low group
un-annotated	2:74436079-74436163	0,000	1755,170	0.0004	0.0319	low group
un-annotated	2:146140303-146140525	0,000	141,576	0.0001	0.0078	low group
un-annotated	3:133864573-133867381	0,000	60,757	0.0001	0.0078	low group
un-annotated	3:144233946-144234106	0,000	190,345	0.0006	0.0370	low group
un-annotated	4:95670907-95689264	0,000	50,452	0.0006	0.0370	low group
un-annotated	4:53167877-53167899	0,000	386984	0.0001	0.0078	low group
un-annotated	6:54625398-54625615	0,000	61,993	0.0005	0.0352	low group

un-annotated	6:71273302-71273742	0,000	14,237	0.0004	0.0308	low group
un-annotated	7:82944504-82947401	0,000	67,349	0.0001	0.0078	low group
un-annotated	7:69929899-69930246	0,000	26,751	0.0001	0.0078	low group
un-annotated	9:133471303-133471700	0,000	14,237	0.0004	0.0319	low group
un-annotated	X:37199874-37200061	0,000	117,755	0.0006	0.0391	low group
CCDC23	4:117723360-117723558	0,000	95,426	0.0003	0.0282	low group
ENSSSCG00000021428	2:137744797-137747670	0,000	34,076	0.0001	0.0078	low group
Gene name	**Locus**	**High RPKM**	**Low RPKM**	***p-value***	***q-value***	**Expression**
un-annotated	1:146326137-146332733	19,218	0,000	0.0001	0.0078	high group
un-annotated	1:153401809-153402129	30,745	0,000	0.0001	0.0078	high group
un-annotated	13:116302834-116303158	70,331	0,000	0.0001	0.0078	high group
un-annotated	15:140262241-140262584	22,064	0,000	0.0004	0.0308	high group
un-annotated	3:100325999-100326309	50,622	0,000	0.0001	0.0078	high group
un-annotated	6:51851923-51854793	14,580	0,000	0.0001	0.0078	high group
ENSSSCG00000009447	11:34047436-34047881	14,081	0,000	0.0001	0.0129	high group
KLK1	6:51469506-51475249	13,053	0,000	0.0001	0.0078	high group
MYL4	12:16799290-16813119	17,438	0,000	0.0001	0.0078	high group
NMU	8:58581166-58601463	152,849	0,000	0.0001	0.0078	high group

Table S3. QTL mapping results for those DEGs located within at least one QTL closely related with litter size. Pocine breeds correspond to: Largewhite (LW), ME (Meishan), Landrance (LD), Duroc (DU), Pietrain (PT), Yorkshire (YS) and French Landrace (French LD).

DEG	Locus	QTL	QTL ID	QTL name	QTL coordinates (bp)	Breed
CES1	6:27276810-27380644	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME
DCLK2	8:83617784-83721850	TNUM	QTL:4253	Teat number	Chr.8:38873367-90653103	
		TNB	QTL:24282	Total number born	Chr.8:3470575-143577862	LW, ME
		OVRATE	QTL:492	Corpus luteum number	Chr.8:52718097-133922781	
ENSSSCG00000000921	5:97776302-97793667	**NSB**	QTL:18128	Number of stillborn	Chr.5:231749-108789684	LW, LD
FXYD3	6:40111306-40118025	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME
GPER	3:828570-833135	**OVRATE**	QTL:4249	Corpus luteum number	Chr.3:14776389-23924094	
		TNUM	QTL:5224	Teat number	Chr.3:1802851-131434519	LW, ME
		DRIPL	QTL:5692	Drip loss	Chr.3:14776389-62780348	DU, PT
		BW	QTL:5694	Body weight (birth)	Chr.3:14776389-62780348	DU, PT
KLK1	6:51469506-51475249	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME
MMP23B	6:58231350-58233984	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME
NEXN	6:125500472-125693958	**TNB**	QTL:10620	Total number born	Chr.6:74531339-129740986	ME, YS
NMU	8:58581166-58601463	TNUM	QTL:4253	Teat number	Chr.8:38873367-90653103	
		TNB	QTL:24282	Total number born	Chr.8:3470575-143577862	LW, ME
		OVRATE	QTL:492	Corpus luteum number	Chr.8:52718097-133922781	
OVOL1	2:5630670-5642984	**21DWT**	QTL:928	Body weight (3 weeks)	Chr.2:2387169-13366532	
PDK4	9:82625076-82638263	**OVRATE**	QTL:517	Corpus luteum number	Chr.9:45173556-138764263	
PION	9:113183480-113208758	**OVRATE**	QTL:517	Corpus luteum number	Chr.9:45173556-138764263	
PTHLH	5:49160138-49172386	**NSB**	QTL:18128	Number of stillborn	Chr.5:231749-108789684	LW, LD
SCNN1G	3:23444853-23480156	**OVRATE**	QTL:4249	Corpus luteum number	Chr.3:14776389-23924094	
		TNUM	QTL:5224	Teat number	Chr.3:1802851-131434519	LW, ME

		BW	QTL:5694	Body weight (birth)	Chr.3:14776389-62780348	DU, PT
		DRIPL	QTL:5692	Drip loss	Chr.3:14776389-62780348	DU, PT
SDCBP2	17:38419446-38437273	WWT	QTL:5231	Body weight (weaning)	Chr.17:3115596-69701581	LW, ME
		TNUM	QTL:5229	Teat number	Chr.17:13961137-69701581	LW, ME
SGPP2	15:137861023-137905123	TNB	QTL:22919	Total number born	Chr.15:134994861-138620895	DU, YS, LD
		NBA	QTL:22930	Total number born alive	Chr.15:134994861-138620895	DU, YS, LD
SH3BGR	13:213286139-213348206	**Wt**	QTL:1139	Body weight (5 weeks)	Chr.13:206615577-218635234	
TM9SF2	11:75703401-75712182	**NSB**	QTL:7534	Number of stillborn	Chr.11:52388584-78227264	French LD, LW
un-annotated	13:3656448-3695027	**NSB**	QTL:18133	Number of stillborn	Chr.13:3477201-3702865	LW, LD
un-annotated	15:117232153-117235744	**OVRATE**	QTL:10614	Corpus luteum number	Chr.15:114074540-153054254	ME, YS
un-annotated	15:122380337-122380596	**OVRATE**	QTL:10614	Corpus luteum number	Chr.15:114074540-153054254	ME, YS
un-annotated	3:100325999-100326309	TNUM	QTL:5224	Teat number	Chr.3:1802851-131434519	LW, ME
		BW	QTL:5234	Body weight (end of test)	Chr.3:2742110-138643006	LW, ME
un-annotated	6:51851923-51854793	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME
un-annotated	3:10903133-10908516	TNUM	QTL:5224	Teat number	Chr.3:1802851-131434519	LW, ME
		OVRATE	QTL:515	Corpus luteum number	Chr.3:2847860-90815870	
un-annotated	6:54625398-54625615	**TNB**	QTL:24281	Total number born	Chr.6:8144130-116389124	LW, ME

Table S4. Differential expressed miRNAs found in small RNA libraries between both extreme phenotypes. All expression values are shown as RPKM values (Reads per Kilobase of exon model per Million mapped reads – Mortazavi et al., 2008). Mean difference between both groups is represented as the $\log_2$ transformed fold change (Log_2FC). Shown *q-values* are Benjamini-Hochberg FDR corrected *p-values.*

miRNA	miRNA precursor	High (RPKM)	Low (RPKM)	Log_2 FC	*p-value*	*q-value*
ssc-let-7c	*ssc-let-7c*	9,495.15	18,854.07	0.99	0.029	0.985
ssc-miR-31	*ssc-mir-31*	56.93	7.46	-2.93	0.004	0.985
ssc-miR-92a	*ssc-mir-92a-2*	51,874.13	21,710.41	-1.26	0.032	0.985
ssc-miR-92a	*ssc-mir-92a-1*	56,610.92	23,545.2	-1.27	0.035	0.985
ssc-miR-101	*ssc-mir-101-1*	430.21	187.19	-1.20	0.034	0.985
ssc-miR-101	*ssc-mir-101-2*	414.3	181.3	-1.19	0.038	0.985
ssc-miR-129a	*ssc-mir-129a*	4,616.22	17,489.47	1.92	0.047	0.985
ssc-miR-144	*ssc-mir-144*	10.64	0.76	-3.81	0.011	0.985
ssc-miR-145-5p	*ssc-mir-145*	34,958.9	81,171.84	1.21	0.045	0.985
ssc-miR-181d-5p	*ssc-mir-181d*	55.51	124.35	1.16	0.046	0.985
ssc-miR-382	*ssc-mir-382*	15.82	45.28	1.52	0.019	0.985
ssc-miR-450c-5p	*ssc-mir-450c*	170.41	411.28	1.27	0.017	0.985

Table S5. TargetScan results showing DEG predicted as target mRNAs for our differentially expressed miRNAs. [a] Sum of the contribution of site-type, 3' pairing, local AU, position, TA (target site abundance) and SPS (seed-pairing stability) calculated as in Garcia et al., 2011 (Garcia et al. 2011). [b] Probability of conserved targeting as described in Friedman et al., 2009 (Friedman et al. 2009)

Representative miRNA	Representative Transcript	Target gene	Conserved sites				Poorly conserved sites				Total context +score[a]	Aggregate P_{CT}[b]
			Total	8mer	7mer-m8	7mer-1A	Total	8mer	7mer-m8	7mer-1A		
hsa-miR-101	NM_053044	*HTRA3*	2	1	0	1	0	0	0	0	-0.49	0.94
	NM_001677	*ATP1B1*	2	0	0	2	0	0	0	0	-0.24	0.52
	NM_000963	*PTGS2*	1	1	0	0	0	0	0	0	-0.22	0.80
	NM_002229	*JUNB*	1	0	1	0	0	0	0	0	-0.06	0.42
hsa-miR-133a	NM_001977	*ENPEP*	1	0	1	0	1	0	0	1	-0.41	0.80
hsa-miR-144	NM_198965	*PTHLH*	1	1	0	0	0	0	0	0	-0.33	< 0.1
	NM_024420	*PLA2G4A*	1	1	0	0	0	0	0	0	-0.30	< 0.1
	NM_001677	*ATP1B1*	2	0	0	2	0	0	0	0	-0.22	0.52
	NM_053044	*HTRA3*	2	0	0	2	0	0	0	0	-0.21	0.52
	NM_000963	*PTGS2*	1	0	0	1	0	0	0	0	-0.06	0.31
hsa-miR-145	NM_001730	*KLF5*	1	1	0	0	0	0	0	0	-0.41	0.47
	NM_021154	*PSAT1*	1	1	0	0	0	0	0	0	-0.30	0.30
	NM_207446	*FAM174B*	1	1	0	0	0	0	0	0	-0.23	0.21
	NM_024420	*PLA2G4A*	1	0	1	0	1	0	0	1	-0.20	0.37
	NM_021914	*CFL2*	1	0	1	0	0	0	0	0	-0.09	0.50
hsa-miR-181d	NM_000902	*MME*	1	0	1	0	1	1	0	0	-0.17	0.35

Table S6. Candidate novel miRNAs predicted by miRDeep in H and L small RNA libraries. Shown RPKM values represent all read counts mapping in each novel miRNA mature, loop and/or star sequence. [a] A miRDeep score of 10 corresponds to a probability of> 90% to be a true positive.

Samples	Coordinates novel miRNA	miRDeep score[a]	Estimated Prob. of True Positive	Total RPKM	Mature RPKM	Loop RPKM	Star RPKM	Significant *p-value*
High	chr5:1-111506441_34692	2.7e+5	0.95 ± 0.04	542,935	542,903	1	31	Yes
	chr6:1-157765593_36025	1.9e+5	0.95 ± 0.04	382,423	382,348	0	75	Yes
	chr13:1-218635234_6943	2.2e+4	0.95 ± 0.04	43,754	43,643	0	111	Yes
	chr5:1-111506441_18208	7.4e+5	0.93 ± 0.06	1,452,784	1,452,741	4	39	Yes
	chr6:1-157765593_19031	3.3e+5	0.93 ± 0.06	652,737	652,542	0	195	Yes
	chr13:1-218635234_3741	1.7e+4	0.93 ± 0.06	33,759	33,661	0	98	Yes
	chr5:1-111506441_23908	9.8e+5	0.94 ± 0.07	1,927,410	1,927,367	0	43	Yes
	chr6:1-157765593_24995	6.7e+5	0.94 ± 0.07	1,324,843	1,324,625	0	218	Yes
	chr13:1-218635234_4737	1.2e+4	0.94 ± 0.07	24,740	24,628	0	112	Yes
	chr6:1-157765593_31265	1.7e+5	0.96 ± 0.04	336,555	336,469	0	86	Yes
	chr5:1-111506441_30154	1.7e+5	0.96 ± 0.04	336,539	336,517	0	22	Yes
	chr13:1-218635234_5831	1.3e+4	0.96 ± 0.04	27,002	26,627	0	375	Yes
	chr5:1-111506441_7732	4.3e+5	0.92 ± 0.06	862,243	862,225	0	18	Yes
	chr6:1-157765593_8073	3.8e+5	0.92 ± 0.06	748,946	748,675	0	271	Yes
	chr13:1-218635234_1581	1.3e+4	0.92 ± 0.06	25,872	25,834	0	38	Yes
Low	chr6:1-157765593_18277	1.1e+5	0.93 ± 0.05	230,090	230,073	0	17	Yes
	chr5:1-111506441_17574	9.8e+4	0.93 ± 0.05	194,035	194,021	0	14	Yes
	chr13:1-218635234_3401	1.4e+4	0.93 ± 0.05	28,734	28,476	0	258	Yes
	chr5:1-111506441_34762	4.2e+5	0.95 ± 0.05	832,924	832,866	0	58	Yes
	chr6:1-157765593_36167	3.7e+5	0.95 ± 0.05	736,278	736,125	0	153	Yes
	chr13:1-218635234_6891	3.3e+4	0.95 ± 0.05	65,306	65,166	0	140	Yes
	chr5:1-111506441_18102	2.1e+5	0.92 ± 0.06	424,999	424,985	0	14	Yes
	chr6:1-157765593_18897	1.3e+5	0.92 ± 0.06	270,022	269,967	0	55	Yes
	chr13:1-218635234_3579	1.7e+4	0.92 ± 0.06	35,288	35,200	0	88	Yes
	chr5:1-111506441_13996	4.0e+5	0.95 ± 0.05	785,564	785,561	0	3	Yes
	chr6:1-157765593_14561	3.1e+5	0.95 ± 0.05	621,594	621,384	0	210	Yes
	chr13:1-218635234_2831	1.4e+4	0.95 ± 0.05	27,973	27,910	0	63	Yes

7.2. **Supplementary material from Study II**: Validation of reproduction-related candidate genes.

Additional file 1. Phenotypic records of the F_2 Iberian × Meishan sows used in this study. [a]NBA and TNB correspond to the average for four consecutive parities. [b]OR and NF were recorded at slaughter on the fifth gestation.

Prolificacy level	Animal	NBA[a]	TNB[a]	OR[b]	NF[b]	EBV
HIGH	A1 (791)	12.00	10.00	13.00	10	1.73
	A2 (787)	11.75	13.00	16.00	16	1.68
	A3 (169)	12.25	11.00	14.00	11	1.68
	A4 (332)	12.75	13.33	16.00	14	1.55
	A5 (373)	11.25	11.00	20.00	17	1.50
	A6 (878)	12.00	10.50	14.00	7	1.42
	A7 (425)	11.00	11.00	0.00	13	1.34
	A8 (767)	9.40	10.50	17.00	14	1.31
	A9 (20)	11.00	10.00	20.00	14	1.22
	A10 (127)	11.00	11.67	17.00	13	1.21
	A11 (365)	10.50	10.00	16.00	9	1.17
	A12 (389)	10.25	10.50	19.00	16	1.09
	A13 (597)	10.00	9.50	20.00	11	0.92
	A14 (151)	10.75	12.00	20.00	13	0.89
	A15 (874	10.25	10.00	11.00	8	0.82
	A16 (271)	10.50	9.67	15.00	14	0.81
	A17 (30)	10.75	10.67	19.00	13	0.80
	A18 (485)	11.00	12.50	16.00	16	0.77
Average (HIGH)		**11.02**	**10.94**	**15.72**	**12.72**	**1.22**
LOW	A19 (350)	4.50	3.00	15.00	6	-2.48
	A20 (309)	5.00	4.33	16.00	8	-2.42
	A21 (360)	5.00	5.33	18.00	1	-2.33
	A22 (260)	4.75	5.00	17.00	10	-2.31
	A23 (173)	5.00	6.67	15.00	10	-2.30
	A24 (861)	5.50	5.00	24.00	9	-2.04
	A25 (409)	4.75	5.67	18.00	11	-1.94
	A26 (918)	7.00	8.50	16.00	13	-1.46
	A27 (779)	6.25	5.50	23.00	10	-1.45
	A28 (915)	4.75	4.00	18.00	8	-1.21
	A29 (443)	5.25	6.50	16.00	5	-1.13
	A30 (702)	6.00	7.50	13.00	11	-1.06
	A31 (322)	4.75	5.00	16.00	14	-0.95
	A32 (204)	5.00	3.67	14.00	15	-0.95
	A33 (486)	5.25	3.50	24.00	5	-0.91
	A34 (499)	6.75	6.50	13.00	11	-0.59
	A35 (895)	7.25	8.50	13.00	10	-0.46
	A36 (846)	6.75	5.00	22.00	14	-0.45
Average (LOW)		**5.53**	**5.51**	**17.28**	**9.50**	**-1.47**

Additional file 2. Predicted and validated expression results for the selected genes.

Gene name	Position	QTL[a]	QuantStudio 12K Flex expression results				RNA-seq expression results					
			H samples (mean RQ)	L samples (mean RQ)	FC	p-value	H samples (RPKM)	L samples (RPKM)	log2 FC	p-value	p-value (FDR)	N
ADM	2:52576214-52578540	NNIP	0.24	0.07	3.34	0.001	165.93	35.82	-2.21	0.0004	0.032	36
CES1	6:27276810-27380644	TNB	0.24	0.07	3.63	0.008	88.82	10.33	-3.10	0.0001	0.008	36
DCLK2	8:83617784-83721850	TNUM TNB OVRATE	0.14	0.23	0.59	0.066	14.48	51.71	1.84	0.0007	0.043	36
FOXA2	17:34053459-34056624	TNUM	0.33	0.35	0.94	0.780	9.49	37.10	1.97	0.0001	0.013	36
FXYD3	6:40111306-40118025	TNB	0.28	0.12	2.41	0.013	211.40	22.67	-3.22	0.0001	0.008	36
GPER[b]	3:828570-833135	OVRATE TNUM DRIPL BW	-	-	-	-	3.93	24.58	2.65	0.0001	0.008	36
IHH	15:134122694-134129391	GEST	0.12	0.24	0.50	0.050	5.93	28.75	2.28	0.0002	0.022	36
KLF5	11:49867087-49870148	TNUM TNUM	0.27	0.08	3.64	0.001	210.13	53.94	-1.96	0.0005	0.036	36
KLK1	6:51469506-51475249	TNB	0.05	0.00	21.33	0.017	13.05	0.00	N/A	0.0001	0.008	36
MMP23B	6:58231350-58233984	TNB	0.13	0.24	0.54	0.073	7.10	55.59	2.97	0.0001	0.008	36
NEXN	6:125500472-125693958	TNB	0.09	0.19	0.49	0.154	11.92	60.29	2.34	0.0001	0.008	36
NMU	8:58581166-58601463	TNUM TNB OVRATE	0.02	0.00	6.80	0.099	152.85	0.00	N/A	0.0001	0.008	36
PDK4	9:82625076-82638263	OVRATE	0.16	0.17	0.96	0.903	22.65	72.26	1.67	0.0005	0.036	36
PION	9:113183480-113208758	OVRATE	0.44	0.27	1.64	0.009	2.92	14.50	2.31	0.0009	0.050	36
SDCBP2	17:38419446-38437273	WWT	0.19	0.09	2.21	0.028	49.49	6.49	-2.93	0.0001	0.008	36

		TNUM										
SGPP2	15:137861023-137905123	TNB NBA	0.05	0.01	4.848	0.125	11.59	1.69	-2.78	0.0001	0.008	36
SH3BGR	13:213286139-213348206	Wt	0.09	0.20	0.448	0.099	14.45	58.62	2.02	0.0001	0.008	36
TM9SF2	11:75703401-75712182	NSB	0.61	0.61	0.992	0.943	65.16	4.27	-3.93	0.0001	0.008	36

[a] QTL identifiers correspond to: Body weight 5 weeks (Wt), Total number born (TNB), Total number born alive (NBA), Body weight at weaning (WWT), Teat number (TNUM), Body weight at 3 weeks (21DWT), Corpus luteum number (OVRATE), Drip loss (DRIPL) , Body weight at birth (BW), Gestation length (GEST), Mummified pigs (MUMM) and Nonfunctional nipples (NNIP). [b] Expression results could not be determined because we could not establish an optimized primer design for the amplification of this gene in our samples.

Additional file 3. Primers used for the genes RT-qPCR validation design.

Gene	Forward primer	Reverse Primer	Type
ADM	GCAGAGTTCCGAAAGAAATGGA	AGGCCCGGCCTTCAAG	Target Gene
B2MG	ACCTTCTGGTCCACACTGAGTTC	GGTCTCGATCCCACTTAACTATCTTG	Endogenous
CES1	AAGTCCTACCCCATCGCTAACA	GTCCCCCCCAAATACTTGTCA	Target Gene
DCLK2	TTTGTACACCGTCTGTGGCAC	TTCAGGCCATAGCCAGTTTCAG	Target Gene
FOXA2	ATGCACTCGGCTTCCAGTATG	TCACCGAGGAGTAGCCCTCG	Target Gene
FXYD3	GGCATCATCATCCTCCTGAGT	TGATCCGTCCTCAACAGTCATG	Target Gene
GPER1[a]	-	-	-
IHH	CTCCGTCAAGTCCGAGCAC	TGACAAGGCCACACGTGC	Target Gene
KLF5	ACGTCTTCCTCCCTGACATCA	GTGGGTTACGCACGGTCTCT	Target Gene
KLK1	AGGACCAGACGACTTCGAATTC	CACAAAACGTATTCTGCAGGAGAGT	Target Gene
MMP23B	TACAGCTGGAAGAAAGGCGTG	GTGGCCGATCTCGTGGG	Target Gene
NEXN	CGGACCTTGGCGTGTTCT	TGGTCGTAGGGTGATTATGAAGCT	Target Gene
NMU	TCCTATTGTAAGCCAAAATCGAAGA	AAATGGGTGGCATTCATTTTAAAT	Target Gene
PDK4	TGCTGGACTTCGGTTCAGAA	GCTAGCCTCACAGGCAACTCTT	Target Gene
PION	AGCTGTCACGAGGCTCATGA	CTGACCGATAAGCGGAGGAA	Target Gene
SDCBP2	GGGCTCCTCACCAACCACTA	GAATCTCTGTGACCTCTTTGTCCTT	Target Gene
SGPP2	CTTGGGACTGGCGTTGGT	CCAGCACGTCCAGGACTGT	Target Gene
SH3BGR	TCTGGGTCCATAGCGATTAGGA	AAAGTCGATTTTATTCGCTTCCA	Target Gene
TM9SF2	CACATTCAGTGGTTTAGCATCATG	ATCATAGCCACCATTCCAGACA	Target Gene
UBC	GCATTGTTGGCGGTTTCG	AGACGCTGTGAAGCCAATCA	Endogenous

[a] Expression level of *GPER1* could not be determined because we could not establish an optimized primer design for the amplification of this gene in our samples.

7.3. **Supplementary material from Study III**: Determination of polymorphisms affecting the regulatory function of reproductive-related miRNAs

Supplementary Table 1. Phenotypic records of the extreme F_2 Iberian × Meishan sows used in this study. [a]NBA and TNB entries correspond to the average for four consecutive parities. and NF recorded at slaughter on the fifth gestation.

Prolificacy level	Animal	NBA[a]	TNB[a]	OR[b]	NF[b]	EBV
HIGH	A1 (791)	12.00	10.00	13.00	10	1.73
	A2 (787)	11.75	13.00	16.00	16	1.68
	A3 (169)	12.25	11.00	14.00	11	1.68
	A4 (332)	12.75	13.33	16.00	14	1.55
	A5 (373)	11.25	11.00	20.00	17	1.50
	A6 (878)	12.00	10.50	14.00	7	1.42
	A7 (425)	11.00	11.00	0.00	13	1.34
	A8 (767)	9.40	10.50	17.00	14	1.31
	A9 (20)	11.00	10.00	20.00	14	1.22
	A10 (127)	11.00	11.67	17.00	13	1.21
	A11 (365)	10.50	10.00	16.00	9	1.17
	A12 (389)	10.25	10.50	19.00	16	1.09
	A13 (597)	10.00	9.50	20.00	11	0.92
	A14 (151)	10.75	12.00	20.00	13	0.89
	A15 (874)	10.25	10.00	11.00	8	0.82
	A16 (271)	10.50	9.67	15.00	14	0.81
	A17 (30)	10.75	10.67	19.00	13	0.80
	A18 (485)	11.00	12.50	16.00	16	0.77
Average (HIGH)		11.02	10.94	15.72	12.72	1.22
LOW	A19 (350)	4.50	3.00	15.00	6	-2.48
	A20 (309)	5.00	4.33	16.00	8	-2.42
	A21 (360)	5.00	5.33	18.00	1	-2.33
	A22 (260)	4.75	5.00	17.00	10	-2.31
	A23 (173)	5.00	6.67	15.00	10	-2.30
	A24 (861)	5.50	5.00	24.00	9	-2.04
	A25 (409)	4.75	5.67	18.00	11	-1.94
	A26 (918)	7.00	8.50	16.00	13	-1.46
	A27 (779)	6.25	5.50	23.00	10	-1.45
	A28 (915)	4.75	4.00	18.00	8	-1.21
	A29 (443)	5.25	6.50	16.00	5	-1.13
	A30 (702)	6.00	7.50	13.00	11	-1.06
	A31 (322)	4.75	5.00	16.00	14	-0.95
	A32 (204)	5.00	3.67	14.00	15	-0.95
	A33 (486)	5.25	3.50	24.00	5	-0.91
	A34 (499)	6.75	6.50	13.00	11	-0.59
	A35 (895)	7.25	8.50	13.00	10	-0.46
	A36 (846)	6.75	5.00	22.00	14	-0.45
Average (LOW)		5.53	5.51	17.28	9.50	-1.47

Supplementary Table 2. Genotypes of the whole population for the identified SNPs. Genotyping was performed from genomic DNA samples of the whole intercross individuals using the KASP™ competitive allele specific PCR genotyping technology.

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
272	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GG
281	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
283	CC	GG	AA	GG	CC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GA
288	TT	CC	GG	AA	TT	GA	CG	TT	AA	TG	GG	GG	GG	AA	AA	GA
291	TC	CG	GA	AG	TC	GG	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
292	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	?	GG	AA	AA	AA
294	CC	GG	AA	GG	CC	AA	CG	TT	AA	?	GG	GG	GG	AA	AA	GG
296	CC	GG	AA	GG	CC	GA	CG	TT	AA	GG	GG	?	GG	AA	AA	GG
299	CC	GG	AA	GG	CC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
488	TC	CG	GA	GG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
490	TC	CG	GA	GG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
585	CC	GG	AA	GG	CC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
601	CC	GG	AA	GG	CC	GG	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
602	TC	CG	GA	AG	TC	GG	GG	TT	AA	TT	GG	?	GG	AA	AA	GA
611	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
624	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
629	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
639	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
645	CC	GG	AA	GG	CC	GG	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
646	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	GG	GG	AA	AA	GA
666	CC	GG	AA	GG	CC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
668	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
675	TT	CC	GG	AG	TT	AA	CG	TT	AA	GG	GG	GA	GG	AA	AA	AA
679	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	GG	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
703	CC	GG	AA	GG	CC	GG	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
710	TT	CC	GG	AA	TT	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
716	TC	?	?	?	?	?	?	TT	AA	?	GG	?	GG	AA	AA	?
719	?	CC	GG	AG	TT	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
721	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
726	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
735	CC	GG	AA	GG	CC	AA	CG	TT	AA	TG	GG	GG	GG	AA	AA	GA
737	TT	CC	GG	AA	TT	GA	GG	CT	GA	TT	GG	?	GG	AA	AA	GG
738	CC	GG	AA	GG	CC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
745	CC	GG	AA	GG	CC	AA	CG	TT	AA	TT	GG	?	GG	AA	AA	GA
746	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
748	TC	CG	GA	AG	TC	GG	CG	TT	AA	TG	GG	AA	GG	AA	AA	GG
755	TC	CG	GA	AG	TC	AA	GG	CT	GA	GG	GG	AA	GG	AA	AA	AA
759	TT	CC	GG	AG	TT	AA	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
774	TT	CC	GG	AA	TT	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GG
775	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GG
776	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GA
777	TC	CG	GA	AG	TC	AA	CC	TT	AA	GG	GG	AA	GG	AA	AA	AA
782	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
784	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	GA	GG	AA	AA	AA
786	TC	CG	GA	AG	TC	AA	CG	CT	GA	GG	GG	GG	GG	AA	AA	GA
787	TT	?	GG	AA	TT	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
791	CC	GG	AA	GG	CC	GA	CG	CT	GA	TT	GG	?	GG	AA	AA	GA
806	TT	CC	GG	AA	TT	GA	CG	TT	AA	GG	GG	GG	GG	AA	AA	GA
807	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
817	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
819	TC	CG	GA	AG	TC	GA	GG	CT	GA	GG	GG	AA	GG	AA	AA	GG
827	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
838	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
857	CC	GG	AA	GG	CC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
859	TC	CG	GA	GG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
876	TT	CC	GG	AA	TT	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
879	CC	GG	AA	GG	CC	GA	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
892	CC	GG	AA	GG	CC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
894	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	GG	GG	AA	AA	AA
900	TT	CC	GG	AA	TT	GA	CG	TT	AA	GG	GG	GG	GG	AA	AA	GA
903	CC	GG	AA	GG	CC	GG	CC	CT	GA	GG	GG	GA	GG	AA	AA	AA
919	TC	CG	GA	GG	TC	AA	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
922	TC	CG	GA	GG	TC	GA	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
925	TC	CG	GA	GG	TC	GA	GG	TT	AA	TT	GG	?	GG	AA	AA	AA
941	CC	GG	AA	GG	CC	GA	GG	TT	AA	TT	GG	?	GG	AA	AA	GA
947	?	?	?	?	?	?	GG	?	AA	GG	GG	?	?	AA	?	?
995	TC	?	?	?	?	?	GG	TT	AA	GG	GG	?	GG	AA	AA	?
1002	TT	CC	GG	AA	TT	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
110808	TT	CC	GG	AG	TT	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
110809	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
110811	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	?	AA	GA
110821	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
110822	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
110825	TT	CC	GG	AA	TT	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
110832	CC	GG	AA	GG	CC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
110849	CC	GG	AA	GG	CC	GG	CC	TT	AA	GG	GG	AA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
110851	TC	CG	GA	AG	TC	GA	CC	CT	GA	GG	GG	AA	GG	AA	AA	GG
110871	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
110876	TT	CC	GG	AA	TT	GA	CC	CT	GA	GG	GG	GG	GG	AA	AA	GA
110877	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	GG	GG	AA	AA	GG
110878	TC	CG	GA	AG	TC	GA	CG	?	GA	GG	GG	GG	GG	AA	AA	GG
110891	TC	CG	GA	GG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
110893	TT	CC	GG	?	TT	?	GG	TT	AA	GG	GG	?	GG	AA	AA	?
110903	CC	GG	AA	GG	CC	GG	CG	CT	GA	TT	GG	?	GG	AA	AA	GG
110906	TT	CC	GG	AG	TT	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GG
110907	TT	CC	GG	AG	TT	?	CG	CT	?	GG	GG	?	GG	AA	AA	?
110908	TT	CC	GG	AG	TT	GG	GG	CT	GA	GG	GG	AA	GG	AA	AA	GG
110916	CC	GG	AA	GG	CC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
110927	TC	?	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
111143	TC	CG	GA	GG	TC	GG	CC	TT	AA	GG	GG	AA	GG	AA	AA	AA
111149	TC	CG	GA	GG	TC	AA	GG	TT	AA	TG	GG	GG	GG	AA	AA	GA
111150	TT	CC	GG	AG	TT	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
111153	TT	CC	GG	AG	TT	AA	GG	TT	AA	GG	GG	?	GG	AA	AA	GA
111155	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
111156	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
111166	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GG	GG	AA	AA	GG
111167	CC	GG	AA	GG	CC	GG	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
111173	TT	CC	GG	AG	TT	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
111174	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	GG	GG	AA	AA	?
111175	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
111343	CC	GG	AA	GG	CC	GA	CG	?	AA	GG	GG	AA	GG	AA	AA	GA
111346	TT	CC	GG	AG	TT	GG	CG	TT	AA	TG	GG	GG	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
111987	CC	GG	AA	GG	CC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GG
111988	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
111990	TT	CC	GG	AA	TT	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
111992	TC	CG	GA	AG	TC	GA	CC	TT	AA	TG	GG	GG	GG	AA	AA	GG
111997	TC	CG	GA	AG	TC	GA	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
200765	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
200766	TT	CC	GG	AG	TT	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
200767	CC	GG	AA	GG	CC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
200777	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GG
200787	TT	?	GG	AG	TT	AA	CG	TT	AA	TT	GG	?	GG	AA	AA	GA
200800	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GG
200829	CC	GG	AA	GG	CC	GA	CG	TT	AA	TT	GG	?	GG	AA	AA	AA
200830	CC	GG	AA	GG	CC	AA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
200837	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
200839	TT	?	GG	AG	TT	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
200850	TC	CG	GA	AG	TC	GG	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
200852	CC	GG	AA	GG	CC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
200861	TC	CG	GA	AG	TC	GG	CG	TT	AA	GG	GG	AA	GG	AA	AA	GG
200862	TT	CC	GG	AA	TT	AA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GG
200865	TT	CC	GG	AA	TT	GG	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
201135	TC	CG	GA	AG	TC	?	CG	CT	GA	TG	GG	AA	GG	AA	AA	?
201144	TC	CG	GA	GG	TC	GG	CG	TT	AA	TG	GG	GG	GG	AA	AA	GA
201145	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
201146	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	GG	GG	AA	AA	GA
201147	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GG
201154	CC	GG	AA	GG	CC	AA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GG

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
201155	CC	GG	AA	GG	CC	AA	CC	TT	AA	TG	GG	AA	GG	AA	AA	GA
201163	CC	GG	AA	GG	CC	AA	GG	CT	?	GG	GG	AA	GG	AA	AA	GA
201166	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
201169	TT	CC	GG	AA	TT	AA	CG	TT	AA	TT	GG	?	GG	AA	AA	AA
201179	TT	CC	GG	AG	TT	AA	GG	TT	AA	TG	GG	GG	GG	AA	AA	AA
201180	TC	CG	GA	AG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
201187	TT	CC	GG	AA	TT	GG	CG	TT	AA	GG	GG	GG	GG	AA	AA	GA
201189	CC	GG	AA	GG	CC	GG	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
201190	TT	CC	GG	AA	TT	AA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
201191	TT	CC	GG	AA	TT	GG	CG	CT	AA	GG	GG	AA	GG	AA	AA	GG
202544	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GG	GG	AA	AA	AA
202545	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GG	GG	AA	AA	GA
202558	TC	CG	GA	GG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
202561	TT	CC	GG	AA	TT	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
202564	CC	GG	AA	GG	CC	AA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
202569	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GG	?	AA	AA	AA
202571	TT	CC	GG	AA	TT	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202573	CC	GG	AA	GG	CC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202576	TT	CC	GG	AA	TT	GG	CC	TT	AA	TG	GG	AA	GG	AA	AA	GA
202580	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
202582	CC	GG	AA	GG	CC	AA	GG	TT	AA	TG	GG	GG	GG	AA	AA	AA
202583	CC	GG	AA	GG	CC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
202584	TC	CG	GA	AG	TC	AA	GG	TT	AA	TG	GG	GG	GG	AA	AA	AA
202585	TT	CC	GG	AA	TT	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
202586	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202591	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
202603	CC	GG	AA	GG	CC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
202604	TC	CG	GA	GG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
202607	TC	CG	GA	GG	TC	GG	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
202609	CC	GG	AA	GG	CC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202612	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
202613	CC	GG	AA	GG	CC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202614	TT	CC	GG	AA	TT	GG	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
202616	TT	CC	GG	AA	TT	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GG
202627	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
202630	CC	GG	AA	GG	CC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
202631	CC	GG	AA	GG	CC	AA	GG	TT	AA	?	GG	GG	GG	AA	AA	AA
202641	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
202642	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
202644	TT	CC	GG	AA	TT	GA	GG	TT	AA	TG	GG	GG	GG	AA	AA	AA
202652	TC	CG	GA	GG	TC	AA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GG
202655	TC	CG	GA	AG	TC	GG	GG	CT	GA	GG	GG	AA	GG	AA	AA	GA
202664	TC	CG	GA	GG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
202674	TC	CG	GA	AG	TC	GG	GG	TT	AA	TT	GG	?	GG	AA	AA	GA
202675	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
202678	TC	CG	GA	GG	TC	GG	GG	TT	AA	TG	GG	AA	GG	AA	AA	GG
202682	TC	CG	GA	GG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
202683	TT	CC	GG	AG	TT	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GG
203727	TT	CC	GG	AA	TT	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GG
203749	TC	CG	GA	AG	TC	GG	CC	TT	AA	GG	GG	GA	GG	AA	AA	AA
203760	TC	CG	GA	GG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
203761	TT	CC	GG	AG	TT	AA	GG	TT	AA	TG	GG	GG	GG	AA	AA	AA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
203773	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
203788	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	AA	GG	AA	AA	GA
203789	TT	CC	GG	AA	TT	AA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
203795	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
203796	CC	GG	AA	GG	CC	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	AA
B790	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
B791	CC	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
B1052	?	GG	AA	GG	CC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
M1	TT	CC	GG	AG	TT	GA	CC	TT	AA	TT	?	?	GG	AA	?	GG
M2	?	CC	GG	AA	TT	GA	GG	TT	AA	?	GG	?	GG	AA	AA	GG
M3	?	CC	GG	AA	TT	GG	CG	CT	GA	TT	?	?	GG	AA	AA	GG
M4	TT	CC	GG	AA	TT	GG	CC	TT	AA	TG	GG	GG	GG	AA	AA	GG
M5	TT	CC	GG	AA	TT	GG	GG	TT	AA	TT	GG	?	GG	AA	AA	GG
M6	TT	CC	GG	AA	TT	GG	CC	CT	GA	?	GG	?	GG	?	AA	GG
M9	TT	CC	GG	AG	TT	GG	CG	TT	AA	TT	?	?	GG	AA	AA	GA
M12	?	CC	GG	AA	TT	GG	CG	CT	GA	GG	GG	GG	GG	AA	AA	GG
M13	?	CC	GG	AA	TT	GG	CG	TT	AA	GG	GG	GG	GG	AA	AA	GA
M14	TT	CC	GG	AA	TT	GG	CG	CT	GA	GG	GG	GG	GG	AA	AA	GA
M15	TT	CC	GG	AA	TT	GG	GG	TT	AA	TT	GG	?	GG	AA	AA	GA
M16	?	CC	GG	AA	TT	GA	CC	CT	AA	TT	GG	?	GG	AA	AA	GG
M17	TT	CC	GG	AA	TT	GG	GG	TT	AA	?	?	?	GG	AA	AA	GA
M19	TT	CC	GG	AG	TT	GG	CG	TT	AA	GG	GG	GG	GG	AA	AA	GG
M20	TT	CC	GG	GG	TT	GG	GG	TT	AA	TT	?	?	GG	AA	AA	GA
M21	TT	CC	?	AG	TT	GA	GG	CT	?	TT	GG	?	GG	AA	AA	GG
MB1	TC	CG	GA	GG	TC	GA	GG	TT	AA	?	GG	AA	GG	AA	AA	GA
MB2	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
MB3	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB4	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB6	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB7	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB8	?	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB9	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	?	AA	GG	AA	AA	GA
MB10	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB11	TC	CG	GA	AG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB12	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB13	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB14	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB15	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB16	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB17	TC	CG	GA	AG	TC	GA	CC	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB18	TC	CG	GA	GG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB19	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB20	TC	CG	GA	GG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB21	?	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB22	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB23	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB24	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB25	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GA
MB26	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GA
MB27	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GA
MB28	TC	CG	GA	AG	TC	GA	?	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB29	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
MB30	TC	CG	GA	AG	TC	GA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB31	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	GA
MB32	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	?	AA	GG	AA	AA	GA
MB33	TC	CG	GA	AG	TC	GA	GG	CT	GA	TG	GG	AA	GG	AA	AA	AA
MB34	TC	CG	GA	AG	TC	AA	GG	CT	?	TG	GG	AA	GG	AA	AA	AA
MB35	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB36	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB37	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB38	TC	CG	GA	AG	TC	AA	GG	CT	GA	GG	GG	GA	GG	AA	AA	AA
MB39	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB40	TC	CG	GA	GG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB41	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB42	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	?	GG	AA	AA	AA
MB43	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB44	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB45	TC	CG	GA	GG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB46	TC	CG	GA	AG	TC	AA	CC	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB47	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB48	TC	CG	GA	AG	TC	GA	CG	CT	AA	GG	GG	AA	GG	AA	AA	GA
MB49	TC	CG	GA	AG	TC	AA	CG	CT	AA	TG	GG	AA	GG	AA	AA	GA
MB50	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB51	TC	CG	GA	AG	TC	AA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB52	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB53	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	?	GG	AA	AA	GA
MB54	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB55	TC	CG	GA	AG	TC	GA	GG	CT	GA	GG	GG	GA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
MB56	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	GA	GG	AA	AA	GA
MB57	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB58	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB59	TC	CG	GA	AG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB61	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB62	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB63	TC	CG	GA	AG	TC	GA	?	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB64	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB65	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB66	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB67	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB68	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB69	TC	CG	GA	AG	TC	AA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB70	TC	CG	GA	AG	TC	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB71	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB72	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB73	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	AA
MB76	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB77	TC	CG	GA	AG	TC	GA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB78	TC	?	GA	AG	TC	GA	CG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB80	TC	CG	GA	AG	TC	AA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB81	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB82	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB83	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB84	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB85	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
MB87	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	AA
MB88	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB89	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB90	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB91	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB92	TC	CG	GA	GG	TC	GA	CC	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB93	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB94	TC	CG	GA	GG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB95	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB96	TC	CG	GA	GG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB97	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB100	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB101	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB102	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB103	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB104	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB105	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB106	TC	CG	GA	AG	TC	GA	CG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB107	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB108	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB109	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	AA	GG	AA	AA	AA
MB110	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB111	TC	CG	GA	AG	TC	GA	CG	CT	GA	GG	GG	GA	GG	AA	AA	AA
MB112	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	AA	GG	AA	AA	GA
MB113	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB114	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA

DNA \ Assay	miR135_1 SNP1	miR135_1 SNP3	miR135_1 SNP4	miR135_1 SNP5	miR135_1 SNP6	miR-27a	miR-106a	miR29b_1 SNP1	miR29b_1 SNP2	miR29b_2 SNP1	miR29b_2 SNP2	miR29b_2 SNP3	miR-195	miR-335	miR-222	miR-146a
MB115	TC	CG	GA	AG	TC	GA	CG	TT	AA	GG	GG	GA	GG	AA	AA	GA
MB116	TC	CG	GA	AG	TC	GA	GG	TT	AA	GG	GG	GA	GG	AA	AA	AA
MB117	TC	CG	GA	GG	TC	AA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB118	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB119	TC	CG	GA	GG	TC	GA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB120	TC	CG	GA	AG	TC	AA	GG	CT	?	TG	GG	AA	GG	AA	AA	GA
MB121	TC	CG	GA	GG	TC	GA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB122	TC	CG	GA	AG	TC	GA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA
MB123	TC	CG	GA	GG	TC	AA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB124	TC	CG	GA	GG	TC	AA	GG	CT	GA	TG	GG	AA	GG	AA	AA	GA
MB125	TC	CG	GA	GG	TC	AA	GG	TT	AA	TG	GG	AA	GG	AA	AA	GA

Supplementary Table 3. Association between the observed genotype for each variant and EBVs. Student's t-test was performed assuming no differences in sample variances (homogeneity of variance or homoscedasticity). Homoscedasticity was estimated by a Levene's statistical test. When a p-value>0.05 was obtained at the Levene's test, no differences in sample variances were assumed. A p-value<0.05 was considered as significant for the T-test.

Variable	Genotype	N	Media	SD	AA vs GG			AA vs AG			AG vs GG		
					p-value	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**
	AA	47	0.35	0.70									
EBV	AG	58	0.22	0.96	0.004	0.628	0.21	0.424	0.131	0.16	0.059	0.498	0.26
	GG	16	-0.28	0.75									

Variable	Genotype	N	Media	SD	AA vs GG			AA vs AG			AG vs GG		
					p-value	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**
	AA	67	0.07	0.95									
EBV	AG	22	0.42	0.61	0.161	-0.268	0.19	0.052	-0.348	0.18	0.680	0.080	0.19
	GG	33	0.33	0.76									

Variable	Genotype	N	Media	SD	GG vs CC			GG vs GC			GC vs CC		
					p-value	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**	***p-value***	**Mean differences**	**SD error**
	GG	63	0.47	0.69									
EBV	GC	44	0.13	0.82	0.0005	1.167	0.27	0.025	0.332	0.15	0.002	0.836	0.26
	CC	15	-0.70	1.00									

Supplementary Table 4. **Mature miRNA expression analysis results**. Expression values of mature miRNAs are shown as mean relative quantities (RQ) and were estimated with qbasePLUS software (Biogazelle) using target specific amplification efficiencies. Relative quantities were normalized for the expression value of uterus reference miRNA *ssc-miR-103.*

Target	p-value	R2	Fold Change		Significant	Genotype	Mean RQ	N
miR-27a	0.006	0.292	AA/AG	1.63	No	AA	0.385	11
			AA/GG	0.52	No	AG	0.236	16
			AG/GG	0.32	Yes	GG	0.739	6
miR-29b-2	1.03E-08	0.798	AA/AG	15.76	Yes	AA	0.257	16
			AA/GG	0.46	No	AG	0.016	5
			AG/GG	0.03	Yes	GG	0.558	5
miR-106a	2.89E-06	0.585	CC/GC	5.81	Yes	CC	0.488	7
			CC/GG	1.34	No	GC	0.084	11
			GC/GG	0.23	Yes	GG	0.365	14

8. AKNOWLEDGMENTS

Aquesta tesi és fruit de la sort.

És fruit de l'enorme sort que he tingut en el meu recorregut fins el moment d'escriure aquestes linies. Per les persones, pels moments viscuts, per l'aprenentatge. Res haguès estat possible sense el recolzament incondicional de tothom que m'ha fet costat en aquesta etapa.

Però si haig de parlar d'incondicionalitat, haig de parlar del Rafa i de la Mari, **els meus pares**. Durant cinc anys de carrera universitària, any i mig de màster i cuatre de tesi s'aprén molt. Moltíssim. Però l'aprenentatge més important de la meva vida no ha estat pas l'acadèmic. Ells, els pares, han estat els millors tutors, directors, professors. La mare hem va donar potser, la primera gran lliçó de la meva vida ensenyant-me qué volia dir "ser persona". Encara ara recordo quant em va costar entedre que la condició de pesona no és inherent a la d' ésser humà. **Mama**, gracias por tu inagotable entrega, por darlo ABSOLUTAMENTE todo para que yo pudiera llegar donde estoy hoy. Por tus sacrificios, por tus noches sin dormir. Por tu comprensión. Por sufrir cada decepción mía como si fuera tuya, por enseñarme el valor de la honestidad y la humildad. Por enseñarme que no hay recompensa sin esfuerzo y que tirar la toalla no es una opción. Por ser un ejemplo de valentía y fuerza. Por ser, y por estar. Te quiero. **Papa**, gracias por convertirme en una todo terreno. Por enseñarme el valor de la tenacidad y el positivismo. Por enseñarme que siempre hay un plan B. Por enseñarme que las princesas también son esas que saben hacerse su corona o tapizarse sus zapatos de cristal. Por entenderme y apoyarme. Te quiero y admiro. A los dos, GRACIAS por ser un referente para mi y convertiros en mi mejor ejemplo.

I ara que parlo d'admiració, em toca mencionar-te a tu **Carlos**, tete. Tu ets la cara més visible d'aquesta tesi, perquè no podía deixar que el teu enomre talent no quedès imprés per sempre en aquest treball. Vuit anys més petit però segles més gran en tot. Estic orgullosa de tu, i l'admiració que m'has transmès ha estat, en molts moments, el motor per fer-me continuar. T'estimo amb bogería, ets la meva debiltat i sempre ho seràs. Tan debó algun cop a la vida pugui aportar a la meva professió la meitat del talent que tu aportaràs a la teva. Gràcies per ser qui en silenci, més ha dit sempre. T'estimo moltísim.

I si de talent va la cosa, és moment de parlar de la teva feina **Armand**. A tú, no per la condició de director sino per la de profesional, van dirigides aquestes linies. Gràcies per la teva confiança (en ocasions quasi cega!) durant aquest camí. Des que vaig començar l'aventura de

la tesi, m'has oferit l'oportunitat de formar-me, d'apendre, de conèixer, d'equivocar-me. D'enfrontar-me als problemes y ensortir-me'n. Bona o no, la científica que avui escriu aquest paràgraf és 90% responsabilitat teva. Gràcies per l'aprenentatge i el creixement profesional i personal que he tingut oportunitat de desenvolupar al teu costat. De tu m'en porto molts bons moments i el saber que de no haver estat pel teu recolzament, no haguès estat capaç de fer ni la meitat del que he estat capaç de fer aquests cuatre anys. El teu suport ha estat imprescindible.

I parlant d'exemples i models científics i personals a seguir, no podía deixar pasar més línies sense mencionar-te **Glòria**. Potser amb una o dues paraules en tindria prou. Potser no em cal ni un paràgraf: T'estimo. El sentiment que millor defineix el que sento per tu es ADMIRACIÓ. Em vas guiar i acompañar en els primers pasos de la meva "vida científica". Em vas donar suport personal i profesional i em vas ensenyar que per ser un bon científic primer s'ha de ser bon amic. Que no és incompatible l'èxit profesional amb l'honestitat. No conec a ningú tan genial a la seva professió que alhora sigui tan humil. Dones lliçons quasi sense parlar Glo, saps que sempre seràs la meva Bossa. Gràcies per obrir-me les portes del 302… em va canviar la vida.

I em van canviar la vida totes les persones darrera aquella porta del 302: Els Marcs, **Culs** i **Maridet**, el **Frantxu**, **l'Estereta**, la **Gemma**, la **Cons**, la **Mertxe** i la **Kine**, la **Palmi**, la **Sara**, la **Mireia**, **l'Oriol**, l'**Ester S**, la **Laura**, l'Ester **Tobi**, l'Ester **Lozi**, el **Siscu**, el **Jordi**… i a tots els que no hi ereu darrera del 302 però heu estat un regal post-màster: **Nekane**, **Irene**….**TOTS**! No sé com de difícil es trobar a tants bons científics junts, però si sé com de difícil es trobar a persones tan especials juntes. A tots vosaltres, per la vostra infinita generositat i amistat: Moltíssimes gràcies. Us admiro enormement.

Però si darrera la porta del 302 vaig trobar persones increíbles, darrera de la porta del Lab. 310 de la tercera planta del CRAG vaig trobar dues amigues absolutament especials. L'Anna, "**la Castelló**", la "mare" de tothom qui comença l'aventura d'esdevenir científic. El que més admiro de tu es la teva discreció. Com ser brillant sense fer soroll. Sense que ni tan sols es noti. Si en algun moment em toca presumir de bona *praxis*, diré que m'ho has ensenyat tu. Sense la teva guía, sense el teu suport, sense les teves lliçons no hauria arribat fins aquí i

sens dubte aquest treball no tindría ni de bon tros l'aspecte que té. Sense tu tot hauria estat molt més complicat, i jo no em sentiría tan orgullosa de la feina feta. Gràcies infinites.

I una cadira més a la dreta la **Betlem**. Tu sempre dius que sóc la nina dels teus ulls. Potser si, però el que si sóc segur és la teva amiga. Als agraïments de les tesis tu sempre protagonitzes els divertits. Però jo no vull agraïr-te només les teves bogeríes i moments de riures infinits, jo et vull donar les gràcies per totes les converses, per totes les vegades que m'has ajudat a mirar endevant i somriure. Per la teva bondat extrema. Perquè ets un tresor, un d'aquests que quan trobes saps que no has de deixar anar. Un producte gourmet amb mala publicitat… que jo no deixaré de consumir mai! (Sé que volies sentir-ho…jaja) Ets una de les persones més especials que he conegut mai. Ets única! I tota paraula amb tu, sobra.

I es clar, com que la cosa ara va de gent autèntica…ha arribat el teu moment, **Vero**. Sé que los Catalanes somos sosos, pero voy a intentar romper ese prejuicio tuyo…aunque espero haberlo roto un poquito antes! A ti podría darte las gracias por muchas cosas, pero voy a dártelas por estar detrás de una puerta al teléfono. Tu bien sabes que aquello fue el principio de todo. Nunca llegué a imaginar el nivel de afinidad que tendría contigo. Y es que nada ha sido nunca predecible contigo. No pierdas jamás esa sinceridad, esa espontaneidad, esa fuerza. Tú me has regalado muchos momentos que no podré olvidar. Gracias por cada uno de ellos.

I es que les portes del CRAG han obert pas a molts moments, vivències i amistats genials. Les primeres, les de la **Ingrid** i **l'Oriol**, companys de grup. Gràcies per transmetre la vostra experiencia a aquella nouvinguda i per haver col·laborat desde la distància en tot allò que he fet després. I ara que parlem de "viejas Glorias", gràcies a vosaltres **Jordi Coromines** i **Xavi**, perquè per a mi sou un clar exemple de que les coses ben fetes tenen una recompensa. Sempre us he vist amb admiració. I per descomptat a tu Anna, la **Merkels!** Gràcies a tú també per ser tan transparent, tan bona persona i tan treballadora. Espero seguir comptant amb el teu exemple molt de temps. A l'**Anna Puig**, la Puchi! Pels petons matiners i les abraçades però sobretot per les hores de converses i el suport a tothora de manera desinteressada. Per fer-me sentir sempre tan recolzada i per tots els consells técnics i no tan técnics a la vora del meu ordinador, "café en mano". Gràcies per la teva amistat, mai oblidaré Helsinki! Y parlar de café sense mecionar a **Monsieur Zidi**, és quasi imperdonable! Zidi, tú sabes de sobras lo

importante que has sido y eres para mí. Mis primeras semanas en el CRAG no habrían sido lo mismo sin tu apoyo. Nuestra amistad que como si de aquellos cafés del principio se tratase, fue surgiendo a fuego lento. Y hasta hoy, que puedo decir que eres de los mejores amigos que me llevo de esta etapa. Gracias por todo. Y recuerda: Siempre nos quedará París!

A l'Spaghetti team: **Eri** (Mrs. Cucharrita!), porque recuerdo nuestros primeros pasos juntas en aquel despacho debatiendo científicamente los misterios religiosos...jaja A ti **Antonia**, porque nunca entenderé como alguien puede dejar tanta huella en tan poco tiempo, y a ti **Ari**, ejemplo de fuerza, tenacidad y responsabilidad. Me has enseñado mucho sin saberlo. A las tres, Grazie per avermi portare energia ed intelligenza in egual misura. A les nuevas y prometedoras generaciones **Johanna**, **Dani**, **Jordi**, **Tainã,** **Rayner**, **Marta,** **Ediane,** **Anna Cuscó,** **Sara** y **Manu.** Estoy convencida de que vais a recolectar muchos éxitos a lo largo de vuestra carrera porque sois unos autenticos CataCRAGs! Me habría encantado tener cuatro años más de tesis para compartir más tiempo con vosotros. A ti **Fabiana**, porque ha sido genial coincidir contigo y poder compartir una intensa semana de charlas nocturnas! Gracias por tu sutileza y tu saber estar. Eres un 10. Te deseo muchísima suerte. Y claro, para personas 10 está **Tania**. ¿Qué seria de nosotros sin tu ayuda? Pero no hablo de papeles, ni de comandas, ni de formularios, ni de plazos… hablo de tu ayuda como amiga. Porque te implicas tanto con todos nosotros que es imposible no dedicar unas líneas a darte las Gracias! Yo que he tenido ocasión de disfrutar de mil desayunos contigo, puedo decir que ojalá siempre seas tan buena con el papeleo como lo eres como persona. Te voy a echar de menos!

I com no mencionar a tres companys que sense voler han estat familia: la **Maria**, el **Jordi Estellé** i el **Yuli**. Als tres i sense distincions: GRÀCIES PER TANT. Per acollir-me, recolzar-me i cuidar-me. Per convertir-vos en una petita familia a la vora del Sena i ensenyar-me el valor de "sentir-se equip". Del magnífic record que guardo dels dies Parisins en sou 100% responsables: em vaig sentir inmensament protegida i estimada. Gràcies per donar-me i ensenyarme tantíssim. Et à vous **Sophie**, de me donner l'occasion d'apprendre de vous. Pour partager votre sagesse et l'expérience. En ouvrant les portes de votre maison et de me faire sentir les bienvenus. Merci beaucoup pour votre amitié et vos précieuses leçons (I hope you excuse my French!). I no puc deixar d'uns altres internacionals, meine lieben Berlinern! **Uwe**, **Antje**, **Dina**, **Samta**, **Hans**, **Neel**, **Nic**, **Lorenzo**, **Alina**, **Scott**, **Mahmoud** und **Michaela**: Ich

danke Ihnen allen! Vielen Dank, dass mein Tag dort so besonders machen. Ich würde nie die Angst vor “schwarzen Bildschirmen” ohne dass Sie verloren haben!

I es clar, el nostre Internacional més Català, l’Àlex (**Mr. Clop**!): Gràcies per les discussions científques i no tan científques i pel teu bon humor. Per les teves paraules d’ànim i per fer-me treure sempre un somriure!

Ara que ja porto cuatre mil pàgines d’agraïments (ja ho saps que sóc molt exagerada, oi **Marcel**?) Em toca parlar de tu. Gràcies per la teva desinteressada ajuda en cada moment d’aquest camí. Per tenir sempre una paraula d’alé i per donar-me tant bons consells. No sé com de fácil és que un becari es faci amic d’un IP, però voilà! Jo t’en considero un de ben bo. Gràcies per ser-hi en tot moment. **Olga**, **Quim**, **Josep María, Miguel**, **Sebas**… vosaltres també sou el perfecte exemple de com es pot ser IP i proper alhora. Tan debó sigui tan bona profesional com vosaltres algun dia.

Per últim no puc oblidar-me dels que fora de la ciència han cuidat de mi desde fa tants anys: els amics de sempre: L’ **“UdG family”** i a vosaltres **“Cercles”**. A TOTS: Gràcies per fer-me sempre costat i omplir els meus dies d’experiències i moments únics. En especial a ti **Raix**. Porque tu eres mi alter ego, mi apoyo incondicional, casi mi hermana. Porque se que eres una de las personas que más se alegra de cualquier éxito que yo alcance y porque no hay nadie más especial que tu. Por ser mi mejor amiga con mayúsculas y estar siempre a mi lado en todo y para todo. Te quiero y lo sabes. Y a ti **Marta**, porque cualquier línea está de más si se trata de darte las gracias. Por último me gustaría dar las gracias a mi familia. La **Córdoba** y la **Terreros**. Porque esta tesis va de eso, de genética, y vosotros habeis contribuido en todo lo que soy hoy. Pero la familia es un concepto que a veces se queda pequeño si se pretende meter en él a personas muy grandes como es el caso de mis tíos **Conchi** y **Pedro** o el de mi tia **Dolores.** Porque si hay alguien en la familia que se sienta tan orgulloso de esto como mis padres esos sois vosotros. Gracias por verme siempre con esos ojos. Os quiero mucho.

Y si empecé estas páginas hablando de incondicionalidad, no podía cerrarlas sin un gracias que se va directo hacia el lugar donde reside todo aquello que la Ciencia aun no ha logrado explorar. A mis iaios, **Carmen** y **Juan**. Porque desde donde estéis, sé que hoy sois las personas más orgullosas del Planeta. Me habría encantado compartir tanta alegría con vosotros. Os quiero, gracias por permitirme crecer con vuestro ejemplo.