SUSANA AMARAL TEIXEIRA

TRANSCRIPTOME ANALYSIS OF PRENATAL DEVELOPMENT IN PIGS

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

Adviser: Simone Eliza Facioni Guimarães

VIÇOSA - MINAS GERAIS
2020

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

T
T226t
2020

Teixeira, Susana Amaral, 1992-
Transcriptome analysis of prenatal development in pigs / Susana Amaral Teixeira. – Viçosa, MG, 2020.
128 f. : il. (algumas color.) ; 29 cm.

Texto em inglês.
Orientador: Simone Eliza Facioni Guimarães.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Suínos - Embrião - Desenvolvimento. 2. RNA-Seq. 3. Organogênese. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-Graduação em Zootecnia. II. Título.

CDD 22. ed. 636.4082

SUSANA AMARAL TEIXEIRA

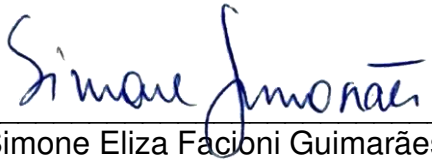
TRANSCRIPTOME ANALYSIS OF PRENATAL DEVELOPMENT IN PIGS

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

APPROVED: October 16, 2020.

Assent:

Susana Amaral Teixeira
Author



Simone Eliza Facioni Guimarães
Adviser

ACKNOWLEDGMENTS

I first would like to thank God for conducted, blessed and allowed me to conclude my studies.

My parents Maria Aparecida and Mozart for all love and example of honesty.

My sister Marina and by brother-in-law Vinicius, for being my support every moment I needed.

My nephews Beatriz and Rodrigo for all love, affection and good moments we had together.

My fiancée Marcos for all love, emotional support, patience and companionship since my under graduation.

All my family, specially my aunts, Çãozinha, Marise and Nivea, for being my example of responsibility and commitment.

My mother-in-law Mônica and my father-in-law Gil Eduardo, for all affection and advices.

My friends, Amanda and Barbara, for all confidence and friendship.

Viçosa city, for allow my personal and professional growth, as well as many good moments since my under graduation.

Universidade Federal de Viçosa and Animal Science Department for support since my under graduation and for all lessons, courses, professionals and good moments offered.

My advisor since under graduation, Professor Simone Eliza Facioni Guimarães, for all attention, incentive, confidence in my work and shared knowledge.

Professors Fabyano Fonseca e Silva, Paulo Savio Lopes and Renata Veroneze for patience, pleasure in helping and advices.

All professors of UFV, in special, Marcio Duarte and Alysson from the Animal Science Department, and José Domingos, from the Veterinary Science Department, for the practical knowledges shared.

Colleagues from GDMA UFV, in special, Arícia, Alessandra, Alex, Delvan, Eula, Hugo, Layla, Leticia, Teo, Talita for the support, knowledge and good moments shared.

My friends from LABTEC Carol, Darlene, Diego, Gustavo, Haniel, Ivan, Ingrid, Larissa, Karine, Margareth, Mariana, Marta, Nathália, Pamela and Walmir for all companionship, good moments and support during these years. In special, Daniele for all friendship and attention, emotional, English writing and statistical support, and Thais for all nice discussions, English writing and funny moments.

Embrapa Suinos e Aves, for allowing me do part of my doctorate degree, in special, Monica Ledur, Jane Peixoto, Adriana Ibelli and Mauricio Cantão for good discussions, shared knowledge and support during my training period at Embrapa.

Dona Inês family', for all affection, friendship and emotional support every time that I needed during my training at Embrapa.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

“Nada do que vivemos tem sentido, se não tocarmos o coração das pessoas” (Cora Coralina)

ABSTRACT

TEIXEIRA, Susana Amaral, D.Sc., Universidade Federal de Viçosa, October, 2020. **Transcriptome analysis of prenatal development in pigs**. Adviser: Simone Eliza Facioni Guimarães.

The conceptuses (embryos and fetuses) developmental trajectory is characterized as a highly regulated biological event at the transcriptional level, which may be influenced by sexual dimorphism and maternal nutrition. Sexual dimorphism at early developmental period, in which the phenotypic sex differentiation is not evident, is mainly related to the expression of genes linked to sex chromosomes. Maternal nutrition is the main extrinsic factor that affects developmental programming, since the molecules from dietary metabolism affects the intrauterine environment, which are important not only to provide nutrients to the conceptuses growth, but also for transcriptional regulation of genes related to placental and conceptuses development. In this context, the effects of pregnant females' supplementation with functional amino acids, such as arginine, have been well studied under productive and reproductive aspects in pig production. However, the influence of this supplementation on global gene expression during conceptuses development is still poorly understood. The proper development and functionality of organs and tissues are events determined during embryogenesis and the fetal period, which require an adjusted control of gene expression. Although in species affected by intrauterine growth restriction, such as pigs, the knowledge about genes expression related to organogenesis can help to understand various aspects associated with prenatal development, studies addressing this event are still scarce. Currently, the gene expression, as well as the biological processes and pathways involved in its regulation, has been performed by transcriptome analysis from RNA sequencing (RNA-Seq) and functional analyses. Thus, in the present study, from the RNA-seq of pig embryos at 25 days of gestation and fetuses at 35 days of gestation, we aimed (1) the conceptuses sex determination; (2) to elucidate

the effects of pregnant gilts supplemented with L-arginine on the conceptuses transcriptome; and, (3) to reveal the main molecular mechanisms involved in pig organogenesis, through the identification of genes differentially expressed between embryos and fetuses. The results of this thesis showed that Y chromosome-linked genes (*DDX3Y*, *KDM5D*, *ZFY*, *EIF2S3Y*, *EIF1AY*, *LOC110255320*, *LOC110257894*, *LOC396706*, *LOC100625207* and *LOC110255257*) presented discrepant reads counts between the embryos and fetuses' samples, which were used for conceptuses sex determination at early prenatal development. Dietary 1.0% L-arginine supplementation of pregnant gilts resulted in downregulation of mitochondrial genes (*ND1*, *ND2*, *CYTB*, *COX2*, *ATP8*, and *mt-rRNA* type) and upregulation of the *CYP1A1* gene in embryos, showing an adjusted transcriptional control at the metabolic level, while the transcriptional program during fetal development was not affected. During the pig organogenesis, 1,705 genes were differentially expressed between embryos and fetuses, from which several biological pathways and processes related to renal, skeletal and bone muscular organogenesis, as well as the development of the cardiovascular and neuronal systems were evidenced. This thesis provided relevant information about the pig prenatal development, which can be applied in studies related to pig production efficiency, as well as greater understanding of developmental biology.

Keywords: Conceptuses. Functional amino acids. Organogenesis. Pigs. Prenatal development. RNA-seq.

RESUMO

TEIXEIRA, Susana Amaral, D.Sc., Universidade Federal de Viçosa, outubro de 2020. **Análise do transcriptoma do desenvolvimento pré-natal de suínos**. Orientadora: Simone Eliza Facioni Guimarães.

A trajetória do desenvolvimento dos conceptos (embriões e fetos) é caracterizada como um evento biológico altamente regulado a nível transcricional, podendo ser influenciada pelo dimorfismo sexual e nutrição materna. O dimorfismo sexual no início do desenvolvimento pré-natal, período no qual a diferenciação sexual com base no fenótipo é dificultada, está principalmente associado à expressão de genes ligados aos cromossomos sexuais. A nutrição materna é o principal fator extrínseco que afeta a programação do desenvolvimento, dada a influência que as moléculas resultantes do metabolismo dietético exercem sobre o ambiente intrauterino, as quais são importantes não apenas por proporcionar nutrientes para o crescimento dos conceptos, mas também por regular a transcrição de genes relacionados ao desenvolvimento da placenta e dos conceptos. Neste contexto, os efeitos da suplementação de fêmeas suínas gestantes com aminoácidos funcionais como a arginina sob aspectos produtivos e reprodutivos têm sido bem estudados. Entretanto, a influência dessa suplementação sobre a expressão gênica global durante o desenvolvimento dos conceptos ainda é pouco conhecida. O adequado desenvolvimento e funcionalidade de órgãos e tecidos são eventos determinados principalmente durante a embriogênese e o período fetal, e requerem um ajustado controle da expressão gênica. Embora em espécies acometidas por restrição do crescimento intrauterino, como os suínos, o conhecimento da expressão de genes relacionados a organogênese possa auxiliar no melhor entendimento de vários aspectos da biologia do desenvolvimento pré-natal este ainda tem sido pouco explorado. Atualmente, o conhecimento da expressão gênica bem como dos processos biológicos e vias envolvidas na regulação gênica tem sido possibilitada pela análise do transcriptoma por meio do sequenciamento de RNA (RNA-Seq) e análises funcionais. Dessa forma, no

presente trabalho, a partir do RNA-seq de embriões aos 25 dias de gestação e fetos suínos aos 35 dias de gestação, nós objetivamos (1) determinar o sexo desses conceptos em estágios iniciais do desenvolvimento; (2) elucidar os efeitos da suplementação de marrãs gestantes com L-arginina sobre o transcriptoma dos conceptos; e, (3) evidenciar os principais mecanismos moleculares envolvidos na organogênese em suínos, por meio da identificação de genes diferencialmente expressos entre embriões e fetos. Os resultados desta tese mostraram que os genes ligados ao cromossomo Y (*DDX3Y*, *KDM5D*, *ZFY*, *EIF2S3Y*, *EIF1AY*, *LOC110255320*, *LOC110257894*, *LOC396706*, *LOC100625207* e *LOC110255257*) apresentaram contagens discrepantes entre as amostras de embriões e fetos suínos as quais foram utilizadas para a determinação do sexo dos conceptos. A suplementação dietética de marrãs com 1% de L-arginina provocou redução na expressão de genes mitocondriais (*ND1*, *ND2*, *CYTB*, *COX2*, *ATP8*, e *mt-rRNA type*) e aumento na expressão do gene *CYP1A1* em embriões, evidenciando um ajustado controle transcricional a nível metabólico, enquanto não provocou alterações transcricionais durante o desenvolvimento fetal. Durante a organogênese em suínos, 1705 genes foram diferencialmente expressos entre embriões e fetos, a partir dos quais várias vias e processos biológicos relacionados à organogênese renal, muscular esquelética e óssea, bem como o desenvolvimento dos sistemas cardiovascular e neuronal, foram evidenciados. Esta tese proporcionou relevantes informações sobre o desenvolvimento pré-natal de suínos que podem ser aplicados em estudos relacionados à eficiência na produção de suínos, bem como no maior entendimento da biologia do desenvolvimento.

Palavras-chave: Aminoácidos funcionais. Conceptos. Desenvolvimento pré-natal. Organogênese. RNA-seq. Suínos.

SUMMARY

CHAPTER 1

General Introduction

Prenatal development is determined by multiples factors, including the intrinsic potential of conceptuses (embryos and fetuses) growth and the availability of nutrients for growth (GODFREY, 1998; PUDDU et al., 2009). The intrinsic potential of growth can be affected by sexual differences, which are determined from sex chromosomes-linked genes at early of development and potentialized by gonad development and hormonal effects, mainly at postnatal development (DEEGAN; ENGEL; ARNOLD, 2019). Regarding nutrients for growth, the maternal nutritional status is important not only to provide nutrients to the conceptuses, but also to modulate the gene expression related to conceptuses and placental development (WU et al., 2004), since the dietary nutrients metabolism is intimately related to the gene expression modulation, mainly by epigenetics modifications (COSTA et al., 2019a; GLUCKMAN et al., 2007; PUDDU et al., 2009; SPINELLI; HAIGIS, 2018; WU et al., 2004).

Among maternal dietary nutrients, arginine is a functional amino acid that has been reported to its positive effects on placental and conceptuses development, mainly by its metabolites, oxide nitric (NO) and polyamines, which are directly involved on placental development, vascularization and blood flow from the mother to the embryo (KESHET; EREZ, 2018; WU et al., 2004, 2009). Due to these versatile functions, the arginine endogenous synthesis does not meet the cellular requirements under certain conditions, such as pregnancy (WU et al., 2018; WU; MORRIS, 1998) and, therefore, it should be provided from diets for pregnant

mammals (LI et al., 2010, 2014; WU et al., 2009). These effects of L-arginine supplementation are mainly important to minimize the consequences of uterine crowding in pregnants sows, such as IUGR (intrauterine growth restriction), since highly-hyperprolific lines have been consistently used in the breed programs (FOXCROFT et al., 2006).

In this context, several studies have reported the effects of maternal supplementation with dietary L-arginine on phenotypic parameters, such as number of surviving pig conceptuses, total number of piglets born and the total weight of the fetuses, which are influenced by supplementation period during pregnancy (for review, see COSTA et al., 2019a). Based on target genes, COSTA *et al*., (2019b) observed that embryos at 25 days of age from L-arginine supplemented gilts during 23 days of gestation showed increased expression of *IGF-1* mRNA, which may result in enhanced embryo growth, since the weight of these embryos also tended to be increased. However, a more complete understanding about the effects of maternal L-arginine supplementation on transcriptional landscape during pig prenatal development is still scarce.

In the prenatal development, the organogenesis is crucial to the establishment of main organs, tissue and body structures, which occurs mainly in embryonic stage (CARDOSO-MOREIRA et al., 2019; HU et al., 2018; HYTTEL et al., 2010). The proper organs and tissues development arise from numerous stimulus resulting of the cell-cell contact and extracellular signals, which are needed to coordinate several biological events, related to cell morphogenesis, proliferation, differentiation and apoptosis. In general, these stimulus act as

upstream signals that are involved in the activation of signaling pathways, initiating a cascade of events that culminate in the activation of transcription factors and, ultimately, the modulating of target genes expressions. Among these pathways, the Hippo and TGFβ/BMP4 signaling pathways are well recognized by their complementary functions (HALDER; JOHNSON, 2011; PAN, 2010), which are crucial to the control of gene expression, mainly related to cell proliferation and differentiation events.

Once the organogenesis occurs, the organs and its associated structure are refined and the whole body grows, characterizing the fetal development (EDWARDS; SAUNDERS; SHIOTA, 2003). During this stage of prenatal development, the specialized pathways related to the cell differentiation and proliferation control, are important to ensure the correct functionally of organ and system (DAGNINO et al., 1997). Among these pathways, genes related to E2F2 transcription factor plays key role required to cell cycle control and differentiation during many biological events, including cardiovascular system and muscle-skeleton development (DAGNINO et al., 1997). Therefore, with the progression of prenatal development, there are mature and complexes organs, resulting in different transcriptional landscape, which means, different set of genes and pathways acting in the spatiotemporal restricted profile (CARDOSO-MOREIRA et al., 2019).

The molecular mechanisms related to prenatal events using pig as model have been elucidated mainly considering embryos or endometrium in the pre-implantation period (LIN et al., 2015; ZENG et al., 2019), since it is a critical period

for prenatal losses (FOXCROFT et al., 2006). On the other hand, during the post-implantation period, which is mainly important in the context of uterine crowding (FOXCROFT et al., 2006), biological events regarding the muscle mass formation at fetal stage has been more evidenced, since in this period the muscle fiber is formed and it is closely related to meat production potential (SOLLERO et al., 2011; TE PAS et al., 2005). However, an overall landscape of molecular events related to pig embryonic and fetal development is still poorly understood.

The knowledge about the main molecular changes related to developmental biology, can be valuable for further studies addressing many aspects involving organogenesis and the overall prenatal development. Additionally, since embryogenesis is marked by high genomic plasticity (HYTTEL et al., 2010), the understanding about molecular changes can be valuable for studies addressing the effects of prenatal insults, such as maternal under or overnutrition, placental flux impairment and teratogens compounds (GODFREY, 1998; PUDDU et al., 2009). In this context, the use of pigs as a model for human prenatal development studies has been employed, since there are numerous genetic and physiological similarities among them (VERMA; RETTENMEIER; SCHMITZ-SPANKE, 2011).

Transcriptome analysis

The molecular basis of phenotypic variation among individuals under two or more conditions (e.g. diseases, nutritional plan, stages of development) has been successfully accessed by a set of differentially expressed (DE) genes, since the messenger RNA (mRNA) is an intermediate molecule between the information contained in the genome and the phenotypic features. Therefore, the identification

of DE genes and its involvement in pathways analyses allow a snapshot of transcriptional activity in different tissues or populations of cells (RITCHIE et al., 2015), in addition to provide valuable information for further studies about the potential molecular mechanisms that modulate it effects (CHEN; LUN; SMYTH, 2016). Due to these biological key roles, the identification and quantification of gene expression has been the core of several studies addressing molecular biology (CONESA et al., 2016).

Among the tools available for identification and quantification of gene transcription, the RNA-Seq (RNA sequencing) tool provides a range of information about the differentially regulated genes through the identification of all groups of genes transcriptionally involved in the modulation of biological functions and signaling pathways, since it is based on high-throughput DNA sequencing (WANG; GERSTEIN; SNYDER, 2009). Moreover, the RNA-seq presents reduced technical variability, that associated with reduced sequencing costs, has been make this tool an attractive method for whole-genome expression studies (VAN VERK et al., 2013). From the broader knowledge of transcriptome, it is possible to access pathways and biological events related to condition under investigation, resulting in a greater support to discuss the molecular mechanisms at transcriptional level (CHEN; LUN; SMYTH, 2016).

The RNA-seq workflow for species in which the genome is well annotated, includes RNA extraction, library preparation and sequencing, raw reads aligning against the genome or transcriptome of the target specie, reads counting, filtering and normalization and, finally the input of these reads for statistical modeling to

obtain the DE genes (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN VERK et al., 2013). These steps are briefly commented bellow.

RNA extraction, library preparation and sequencing

The RNA-seq uses the next-generation sequencing (NGS) methods to sequence the complementary DNA molecule (cDNA) (VAN VERK et al., 2013), in which the first step includes the total RNA extraction from a subpopulation of cell or tissue under evaluated experimental conditions (DÜNDAR; SKRABANEK; ZUMBO, 2018). In this step, the research interest should be considered based on mRNA or small RNA molecules evaluation, since the extraction protocols is specific for each case (DÜNDAR; SKRABANEK; ZUMBO, 2018). Most commonly, the extraction of the total mRNA is performed, followed by an enrichment protocol in order to increase the mRNA abundance, which is carried out by poly(A)-enrichment or rRNA depletion (CONESA et al., 2016; DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019). The poly(A) selection is indicated when the mRNA abundancy is not a limiting factor and a minimal degradation of the extracted mRNA is observed, which can be determined from the RNA integrity number (RIN) value of at least eight (CONESA et al., 2016; VAN VERK et al., 2013).

Afterwards, the high-quality mRNA is chemically (e.g. metal ion exposure), enzymatically (e.g., RNAses) or physically (e.g., shearing) fragmented in a range of 150 – 300 base pairs (bp) of length, and a reverse transcription into cDNA is performed (VAN DEN BERGE et al., 2019; DÜNDAR; SKRABANEK; ZUMBO, 2018). At this moment, adaptors sequences are inserted in the cDNA molecules 3’

and 5’ extremities, which are used as primers in the amplification phase by polymerase chain reaction (PCR) (VAN DEN BERGE et al., 2019). Numerous sequencing methods are available, including the sequencing by synthesis methodology of the Illumina platform (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019).

Briefly, this sequencing protocol is based on flow cell loading with cDNA fragments, which form clusters of double-stranded DNA from hybridization of the adapter sequences with complementary short oligonucleotide of the flow cell, producing several clusters of fragments that are massively and clonally sequenced (DÜNDAR; SKRABANEK; ZUMBO, 2018). In this platform, the sequencing can be performed by the single-end (SE) protocol, in which only one extremity of the cDNA insert is sequenced, or paired-end (PE) protocols, in which both extremities are sequenced, resulting in two reads of the opposite orientation, one from each extremity (VAN DEN BERGE et al., 2019). The choice of the best protocol for library sequencing preparation should be guided by the aims of the analysis and knowledge of the genome of the species studied (CONESA et al., 2016). Furthermore, in the sequencing protocols, the information about the strand from which the read is originated can be preserved (stranded) or not (unstranded). The stranded method can be advantageous, since the expressed RNAs can originate from opposite strands of the same genomic locus, while in the unstranded methodology, the analysis and quantification of antisense or overlapping transcripts is impaired (CONESA et al., 2016; VAN DEN BERGE et al., 2019).

Finally, each nucleotide base of each sequenced fragment emits a fluorescent signal from fluorescently labeled deoxynucleoside triphosphate (dNTP) that is added in each reaction (DÜNDAR; SKRABANEK; ZUMBO, 2018 VAN DEN BERGE et al., 2019). This fluorescent signal is recorded and, based on the excitation spectra, the bases are inferred and ordered (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019). In the sequencing library, each of these sequenced fragments are known as “reads”, which are produced in the range of the millions and are used as an accurate measure of the gene transcription levels in the further RNA-seq steps (VAN VERK et al., 2013).

Mapping, counting, filtering and normalization of reads

After the sequencing, the read alignment or mapping is a crucial step to designate each read to its more probable genomic region of origin (DÜNDAR; SKRABANEK; ZUMBO, 2018). The quality of mapping is a global indicator of the overall sequencing accuracy and of the presence of contaminating DNA, which is measured by the percentage of mapped reads, being 70 to 90% a range considered as a regular RNA-seq reads to map onto the human genome, for instance (CONESA et al., 2016). In order to improve the mapping rate, it is recommended to performed the trimming of the raw reads before to alignment step (BOLGER; LOHSE; USADEL, 2014; CONESA et al., 2016). This trimming is based on cutoff parameters that include low-quality reads, adaptor sequences, and poor-quality of the nucleotide bases (BOLGER; LOHSE; USADEL, 2014; CONESA et al., 2016). Among trimming tools, the Trimmomatic software has been widely used,

given its flexibility to handle PE data and high performance, which are maximized for Illumina sequencing data (BOLGER; LOHSE; USADEL, 2014).

Considering the structural complexity of the eukaryote genome, some reads can be derived from cDNA molecules that correspond to a fully (or partially) spliced transcripts, resulting in the exon-exon-spanning reads that can be separated by tens of thousands of nucleotides (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019). The exon-exon-spanning reads alignment is considered the main challenge of RNA-seq data, since it can lead to inappropriate discard of the reads by multimapping (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019). In order to minimize this, the use of the splice-aware aligners is recommended, such as the HISAT2 (Hierarchical Indexing for Spliced Alignment of Transcripts) and STAR (Spliced Transcripts Alignment to a Reference) algners that perform the spliced alignment against a reference genome (DOBIN et al., 2013; KIM; LANGMEAD; SALZBERG, 2015; VAN DEN BERGE et al., 2019).

From the alignment, the number of reads that were mapped within a genomic feature of interest (such as a gene or an exon) in the chromosomes, which is further used as a measure of the feature abundance in the analyzed sample (SONESON; DELORENZI, 2013), is accessed by software, such as HTSeq (high-throughput sequencing) (SRINIVASAN; VIRDEE; MCARTHUR, 2019) and featureCounts (LIAO; SMYTH; SHI, 2014). In the HTSeq, the reads count within exon or genes is obtained from a file containing genomic coordinates, which are values attributed to genomic positions (e.g. read coverage) or genomic intervals (e.g. exons or genes) (ANDERS; PYL; HUBER, 2015).

However, before directly compare the expression levels among the samples, the reads count need to be correctly filtered, in order to remove the consistently low counts among samples, since the low counts do not provide enough statistical evidence for a reliable biological interpretation of the outcomes and can alter the number of statistical tests (CHEN; LUN; SMYTH, 2016). Moreover, these reads need to be normalized to ensure the similarity among the expression distributions in each sample (SMYTH et al., 2016; VAN DEN BERGE et al., 2019). In general, normalization methods based on library size scaling factor for each sample is the most appropriate, such as the TMM (trimmed mean of M-values, where M-values refers empirical fold changes between two samples) implemented in DE analysis programs, such as edgeR and limma (DÜNDAR; SKRABANEK; ZUMBO, 2018; VAN DEN BERGE et al., 2019).

Differentially Expressed Genes Analysis

In order to obtain the genes that are differentially regulated under investigated conditions, the RNA-seq experimental design requires sufficient numbers of replicates, which is influenced by the technical and biological variability (DÜNDAR; SKRABANEK; ZUMBO, 2018). The technical variability is the result of RNA-seq procedures, such as library preparation and distribution of samples in the sequencing lanes, while the biological variability is influenced by the heterogeneity of the experimental samples, such as gender and bacterial colonies (CONESA et al., 2016). Whereas the technical variation is lower and well controlled from RNA-seq procedures well stablished for all samples, the biological variation is inherent to each individual and can greatly influence on the proper gene transcription

(HANSEN et al., 2011; DÜNDAR; SKRABANEK; ZUMBO, 2018). Therefore, minimizing the variations sources and establishing an adequate number of replicates, where a minimum of three biological replicates is recommended, in order to increase the statistical power, and detect DE genes that are statistically significant among the experimental groups of interest (CONESA et al., 2016).

The first point in the DE analysis consists in summarizing the normalized reads for each feature (e.g. genes) into a counting matrix, which contain the reads counting that remained after filtering in each row, and the samples identification in each column (DÜNDAR; SKRABANEK; ZUMBO, 2018; RITCHIE et al., 2015; VAN DEN BERGE et al., 2019). This file is submitted to statistical modeling, which is based on multiple hypothesis tests for each gene DE among the samples, followed by a *p-value* adjust for multiple testing, in order to avoid an increased rate of false positives results (VAN DEN BERGE et al., 2019). This correction is commonly performed using Benjamini & Hochberg methodology (BENJAMINI; HOCHBERG, 1995).

With the decrease in DNA sequencing costs, the number of studies adressing DE analysis from RNA-seq has increased, resulting in the increase of data to statistical analyses as well as in the increases in the complexity of the experimental design (CUI et al., 2016). However, the DE analysis from the high throughput data as RNA-seq technology is relatively new, and numerous methodologies are being continuously developed (SONESON; DELORENZI, 2013), which leads to lack of consensus about the best approach or software for these analyses (CONESA et al., 2016; ZHAO et al., 2016). Among the popular

softwares, the edgeR and DESeq, involve count-based models and negative binomial (NB) distribution (CONESA et al., 2016; LAW et al., 2014; VAN DEN BERGE et al., 2019).

Recently, the use of the NB distribution for accurate DE analyses has been questioned in the context of various steps of data normalization, including TMM, which can change the nature of distribution from discrete to continuous (CONESA et al., 2016). In addition, the currently available count-based distributions methods limit statistical analysis of RNA-seq data from experiments containing random effects, including repeated measures on the same individual or RNA samples collected at the same time (LAW et al., 2014). In this context, methods based on reads counts transformation, such as the methods performed by limma, has shown reliable results even under these challenging conditions, associated with greater speed of RNA-seq data processing (LAW et al., 2014; RITCHIE et al., 2015).

The limma methodology was firstly developed for microarray data and has been successfully implemented for RNA-seq data from minimal changes in the original pipeline, allowing differential expression analysis from complex experimental designs and testing flexible hypotheses, besides to allow reliable statistical conclusion even when the number of samples is limited (RITCHIE et al., 2015). The general procedures to these analyses from limma, include the convertion of the normalized counting reads into logarithmic scale, which are fitted to linear models, and estimation of mean-variance relationship robustly and non-parametrically by empirical Bayes statistical methods (LAW, 2014; RITCHIE et al., 2015). The mean-variance relationship can be incorporated from a empirical Bayes

procedure (limma-trend) or from inclusion of the mean-variance trend information into a precision weight (limma-voom) (LAW et al., 2014; RITCHIE et al., 2015). The limma trend approach is recommended when the ratio of the largest to the smallest library size is less than three, while the limma-voom performance has been shown superior in cases where the library size between samples is more different (SMYTH; RITCHIE; THORNE, 2019; LAW et al., 2014).

Once defined the methdology for acesses the DE genes, the high-throughput transcriptome frequently generates a long list of genes, from which the main goal is the functional characterization of these genes, including molecular functions and pathways that are related to these genes, and their association with the experimental condition under investigation (CONESA et al., 2016; HUANG; SHERMAN; LEMPICKI, 2009a). This is one of the most challenging efforts in a typical transcriptome study, since there is no gold standart that defines how these long list should be analysed (HUANG; SHERMAN; LEMPICKI, 2009a; VAN DEN BERG et al., 2009). A simple and widely used strategy is to perform a grouping of genes based on their ontology, highlighting the biological processes, molecular function and cellular components that are overrepresentated (YOUNG et al., 2010), as discussed below.

Functional Analyses

Since 2002, several approaches have been developed to analyse the ontological terms related to differentially regulated genes, mainly from microarray studies, including DAVID, BiNGO, FatiGo and GOstats functional tools (HUANG; SHERMAN; LEMPICKI, 2009). Among these functional tools, DAVID (Database for

Annotation, Visualization and Integrated Discovered) is recognized as an web-based database able to handle to any types of gene list from differents genomics platforms, including microarray and RNA-seq data, and it uses the most traditional algorithm to calculate the statistical enrichment of the gene from a set of pre-selected genes by the user, such as the DE genes at determined *p-value* and fold change (HUANG; SHERMAN; LEMPICKI, 2009b).

On DAVID, the users provide a gene list containing the gene identifier (ID) of the all sigificative DE genes, which is used as input list, and define a background list, a reference list from which the degree of gene enrichment in each functional annotation category is defined (HUANG et al., 2009). As result, the enriched functional categories (biological processes, cellular components and pathways) are discussed at level of a genes set, instead of an individual gene, which allows the proper identification of the most important biological events related to the study condition (HUANG; SHERMAN; LEMPICKI, 2009b). In order to guide the disscusion about the functional categories enriched by a set of genes under investigation, some parameters are provided by DAVID, including enrichment score, fold enrichment, *p-value* and corrected *p-value* for multiple testing, such as Benjamini and Bonferroni methods, which instead of establish a rule for the best analyse, emphasize a critical evaluation of the users regarding the biological relevance of the DAVID outcomes (HUANG; SHERMAN; LEMPICKI, 2009a).

As a consequence of the popularity of high-throughput techniques, more genes may be detected and thus, more gene ontologies can be released (SUPEK et al., 2011). This often results in a larger size functional annotation list, which

contains several similar GO terms, impairing the biological interpretation of the results. In order to mitigate these redundant similarities and facilitate the summarizing and visualizing of the relevant GO terms, some tools based on “semantic similarity” terms were proposed, including REVIGO (SUPEK et al., 2011). Basically, from the output enriched GO terms provided by an enrichment-based software, REVIGO forms clusters of the GO similar terms based on “semantic similarity” and displays only the representatives GO, which can be visualized and uploaded in several formats, including two-dimensional plots, interactive graphs and tree map (SUPEK et al., 2011).

In the context of pathway databases, the Toppcluster is another usual tool to performer functional analyses. This web-based tool, in addition to perform gene ontology enrichment, allows a broader pathway enrichment analyses by integration of several pathways annotation databases, including KEGG, REACTOME, PantherDB and Pathway Interaction Database (https://toppgene.cchmc.org/navigation/database.jsp). Besides that, the ToppCluster yields a rich functional map as output that allow the visualization of specifics and shared list of functional features (KAIMAL et al., 2010).

This thesis: objectives

This thesis describes the transcriptional changes during pig conceptuses development from RNA-seq analysis. In Chapter 2, we proposed the conceptuses sex determination by reads counting mapped on chromosome Y-linked genes, since at early prenatal development, the gender is not phenotypically divergent. In Chapter 3, the transcriptional changes during embryos and fetuses’ development

from gilts dietary L-arginine supplemented during gestation were evaluated. Moreover, gene ontology network based on biological processes enriched for DE genes were performed to better understanding and visualization of the effects of maternal supplementation with dietary L-arginine on pig prenatal development. In Chapter 4, we unravel the main biological processes and pathways involved during embryonic to fetal transition, from the DE genes between the proposed developmental stages in pigs. Finally, in Chapter 5 an overall conclusion regarding the use of RNA-seq approach for transcriptome analysis of pig prenatal development was reported, highlighting the main outcomes obtained in this thesis.

References

ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq-A Python framework to work with high-throughput sequencing data. **Bioinformatics**, v. 31, n. 2, p. 166–169, 2015.

BENJAMINI, Y.; HOCHBERG, Y. Controllin the false discovery rate: a practical and powerful approch to multiple testing. **Journal of the Royal Statistic Society**, v. 57, p. 289–300, 1995.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.

CARDOSO-MOREIRA, M. et al. Gene expression across mammalian organ development. **Nature**, 2019.

CHEN, Y.; LUN, A. T. L.; SMYTH, G. K. From reads to genes to pathways: Differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline [version 1; referees: 5 approved]. **F1000Research**, v. 5, p. 1–48, 2016.

CONESA, A. et al. A survey of best practices for RNA-seq data analysis. **Genome Biology**, v. 17, p. 1–19, 2016.

COSTA, K. A. et al. Nutrition influence on sow reproductive performance and conceptuses development and survival: A review about L -arginine supplementation. **Livestock Science**, v. 228, n. July, p. 97–103, 2019a.

COSTA, K. A. et al. Dietary L-arginine supplementation during early gestation of gilts affects conceptuses development. **Theriogenology**, v. 140, p. 62–71, 2019b.

CUI, S. et al. What if we ignore the random effects when analyzing RNA-seq data in a multifactor experiment. **Statistical Applications in Genetics and Molecular Biology**, v. 15, n. 2, p. 87–105, 2016.

DAGNINO, L. et al. Expression patterns of the E2F family of transcription factors during murine epithelial development. **Cell Growth and Differentiation**, v. 8, n. 5,

p. 553–563, 1997.
DEEGAN, D. F.; ENGEL, N.; ARNOLD, A. P. Sexual Dimorphism in the Age of Genomics : How , When , Where. v. 7, n. September, p. 1–7, 2019.
DOBIN, A. et al. STAR: Ultrafast universal RNA-seq aligner. **Bioinformatics**, v. 29, n. 1, p. 15–21, 2013.
DÜNDAR, F.; SKRABANEK, L.; ZUMBO, P. Introduction to differential gene expression analysis using RNA-seq. **Applied Bioinformatics Core**, n. September 2015, p. 1–89, 2018.
EDWARDS, M. J.; SAUNDERS, R. D.; SHIOTA, K. Effects of heat on embryos and foetuses. **International Journal of Hyperthermia**, v. 19, n. 3, p. 295–324, 2003.
FOXCROFT, G. R. et al. The biological basis for prenatal programming of postnatal performance in pigs. **Journal of animal science**, v. 84 Suppl, p. 105–112, 2006.
GLUCKMAN, P. D. et al. Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 31, p. 12796–12800, 2007.
GODFREY, K. M. Maternal regulation of fetal development and health in adult life. **Obstetrics & Gynecology**, v. 78, p. 141–150, 1998.
HALDER, G.; JOHNSON, R. L. Hippo signaling: Growth control and beyond. **Development**, v. 138, n. 1, p. 9–22, 2011.
HANSEN, K. D. et al. Sequencing technology does not eliminate biological variability. **Nature Biotechnology**, v. 29, n. 7, p. 572–573, 2011.
HU, B. et al. Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. **Genome Biology**, v. 19, n. 1, p. 1–20, 2018.
HUANG, D. W. et al. DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 169–175, 2009.
HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature Protocols**, v. 4, n. 1, p. 44–57, 2009a.
HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. **Nucleic Acids Research**, v. 37, n. 1, p. 1–13, 2009b.
HYTTEL, P. et al. **Essentials of domestic animal embriology**. [s.l.] Elsevier, 2010.
KAIMAL, V. et al. ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and networkbased dissection of biological systems. **Nucleic Acids Research**, v. 38, n. SUPPL. 2, p. 96–102, 2010.
KESHET, R.; EREZ, A. Arginine and the metabolic regulation of nitric oxide synthesis in cancer. **Disease Models & Mechanisms**, v. 11, n. 8, p. dmm033332, 2018.
KIM, D.; LANGMEAD, B.; SALZBERG, S. L. HISAT: a fast spliced aligner with low memory requirements. **Nat Methods**, v. 12, n. 4, p. 357–360, 2015.
LAW, C. W. et al. voom : precision weights unlock linear model analysis tools for

RNA-seq read counts. p. 1–17, 2014.
LI, X. et al. Dietary Supplementation with 0.8% L-Arginine between Days 0 and 25 of Gestation Reduces Litter Size in Gilts. **The Journal of Nutrition**, v. 140, n. 6, p. 1111–1116, 2010.
LI, X. et al. Dietary supplementation with l-arginine between days 14 and 25 of gestation enhances embryonic development and survival in gilts. **Amino Acids**, v. 46, n. 2, p. 375–384, 2014.
LIAO, Y.; SMYTH, G. K.; SHI, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. **Bioinformatics**, v. 30, n. 7, p. 923–930, 2014.
LIN, H. et al. Transcriptomic Analysis of the Porcine Endometrium during Embryo Implantation. **Genes**, p. 1330–1346, 2015.
PAN, D. The hippo signaling pathway in development and cancer. **Developmental Cell**, v. 19, n. 4, p. 491–505, 2010.
PUDDU, M. et al. The kidney from prenatal to adult life: Perinatal programming and reduction of number of nephrons during development. **American Journal of Nephrology**, v. 30, n. 2, p. 162–170, 2009.
RITCHIE, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. **Nucleic acids research**, v. 43, n. 7, p. 1–13, 2015.
SMYTH, G. K. et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. **F1000Research**, v. 5, p. 1–30, 2016.
SMYTH, G. K.; RITCHIE, M.; THORNE, N. Linear Models for Microarray and RNA-seq Data User ' s Guide. **Bioinformatics**, p. 1–148, 2019.
SOLLERO, B. P. et al. Transcriptional profiling during foetal skeletal muscle development of Piau and Yorkshire-Landrace cross-bred pigs. **Animal Genetics**, v. 42, n. 6, p. 600–612, 2011.
SONESON, C.; DELORENZI, M. A comparison of methods for differential expression analysis of RNA-seq data. **BMC Bioinformatics**, v. 14, 2013.
SPINELLI, J. B.; HAIGIS, M. C. The multifaceted contributions of mitochondria to cellular metabolism. **Nature Cell Biology**, v. 20, n. 7, p. 745–754, 2018.
SRINIVASAN, K. A.; VIRDEE, S. K.; MCARTHUR, A. G. Strandedness during cDNA synthesis , the stranded parameter in htseq-count , and analysis of RNA-Seq data. n. March, p. 10–11, 2019.
SUPEK, F. et al. Revigo summarizes and visualizes long lists of gene ontology terms. **PLoS ONE**, v. 6, n. 7, 2011.
TE PAS, M. F. W. et al. Transcriptome expression profiles in prenatal pigs in relation to myogenesis. **Journal of Muscle Research and Cell Motility**, v. 26, n. 2, p. 157–165, 2005.
VAN DEN BERG, B. H. J. et al. Comparing gene annotation enrichment tools for functional modeling of agricultural microarray data. **BMC Bioinformatics**, v. 10, n. SUPPL. 11, p. 1–9, 2009.
VAN DEN BERGE, K. et al. RNA Sequencing Data: Hitchhiker's Guide to Expression Analysis. **Annual Review of Biomedical Data Science**, v. 2, n. 1, p. 139–173, 2019.
VAN VERK, M. C. et al. **RNA-Seq: Revelation of the messengersTrends in**

Plant Science, 2013.
VERMA, N.; RETTENMEIER, A. W.; SCHMITZ-SPANKE, S. Recent advances in the use of Sus scrofa (pig) as a model system for proteomic studies. **Proteomics**, v. 11, n. 4, p. 776–793, 2011.
WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, p. 57–63, 2009.
WU, G. et al. Board-invited review: Arginine nutrition and metabolism in growing, gestating, and lactating swine. **Journal of Animal Science**, v. 96, n. 12, p. 5035–5051, 2018.
WU, G. et al. Arginine metabolism and nutrition in growth, health and disease. **Amino Acids**, v. 37, n. 1, p. 153–168, 2009.
WU, G. et al. Recent Advances in Nutritional Sciences Maternal Nutrition and Fetal. **The Journal of Nutririon**, n. 13, p. 2169–2172, 2004.
WU, G.; MORRIS, S. M. Arginine metabolism: nitric oxide and beyond. **The Biochemical journal**, v. 336, p. 1–17, 1998.
YOUNG, M. D. et al. Gene ontology analysis for RNA-seq: accounting for selection bias. **Genome Biology**, v. 11, n. 2, 2010.
ZENG, S. et al. Differential transcriptome dynamics during the onset of conceptus elongation and between female and male porcine embryos. **BMC Genomics**, v. 20, n. 1, p. 1–19, 2019.

CHAPTER 2

Sex Determination Using RNA-Sequencing Analyses in Early Prenatal Pig Development

Susana A. Teixeira[1], Adriana M. G. Ibelli[2], Maurício E. Cantão[2], Haniel C. de Oliveira[1], Mônica C. Ledur[2], Jane de O. Peixoto[2], Daniele B.D. Marques[1], Karine A. Costa[1], Luiz. L. Coutinho[3] and Simone E. F. Guimarães[1,*]

[1]Animal Science Department, Universidade Federal de Viçosa, Viçosa 36570-900, Brazil; susana.amaral.teixeira@gmail.com (S.A.T.); hanielcedraz@gmail.com (H.C.O.); danielebdiniz@gmail.com (D.B.D.M.); kryneacosta@yahoo.com.br (K.A.C.)

[2]Animal Genetics Laboratory, Embrapa Swine and Poultry National Research Center, Concórdia 89715-899, Brazil; adriana.ibelli@embrapa.br (A.M.G.I.); mauricio.cantao@embrapa.br (M.E.C); monica.ledur@embrapa.br (M.C.L.); jane.peixoto@embrapa.br (J.O.P.)

[3]Functional Genomics Center, ESALQ, Universidade de São Paulo, Piracicaba 89715-899, Brazil; llcoutinho@usp.br (L.L.C.)

*Correspondence: sfacioni@ufv.br; Tel.: +55 31 36124625.

Genes - MDPI
Received: 30 October 2019;
Accepted: 27 November 2019;
Published: 5 December 2019

Abstract: Sexual dimorphism is a relevant factor in animal science, since it can affect the gene expression of economically important traits. Eventually, the interest in the prenatal phase in a transcriptome study may not comprise the period of development in which male and female conceptuses are phenotypically divergent. Therefore, it would be interesting if sex differentiation could be performed using transcriptome data, with no need for extra techniques. In this study, the sex of pig conceptuses (embryos at 25 days-old and fetuses at 35 days-old) was determined by reads counts per million (CPM) of Y chromosome-linked genes that were discrepant among samples. Thus, ten genes were used: *DDX3Y*, *KDM5D*, *ZFY*, *EIF2S3Y*, *EIF1AY*, *LOC110255320*, *LOC110257894, LOC396706, LOC100625207,* and *LOC110255257*. Conceptuses that presented reads CPM sum for these genes (ΣCPM_{chrY}) greater than 400 were classified as males and those with ΣCPM_{chrY} below 2 were classified as females. It was demonstrated that the sex identification can be performed at early stages of pig development from RNA-sequencing analysis of genes mapped on Y chromosome. Additionally, these results reinforce that sex determination is a mechanism conserved across mammals, highlighting the importance of using pigs as an animal model to study sex determination during human prenatal development.

Keywords: early gestation; conceptuses development; conceptuses sex determination; pig transcriptome

1. Introduction

Maternal nutrition may affect progeny anatomy, physiology and metabolism during critical periods of prenatal life, which is described as fetal programming [1].

The effects of maternal nutrition on fetal programming have been recently addressed in several studies using pigs as a model organism [2–6], since there are striking similarities between pigs and humans regarding the anatomy, physiology, metabolism, and nutrition, which provide the basis for the use of this animal in several studies. In addition, Sus scrofa is a prolific specie with a relatively short developmental period, besides being a cheaper and easier mammalian model organism to maintain [7,8]. In this context, transcriptome studies using RNA-sequencing (RNA-seq) performed in pigs highlight the relevance of the data obtained in this specie for human-related researches.

The RNA sequencing (RNA-seq) approach generates a comprehensive picture of gene expression levels at different developmental stages and physiological conditions [9–11]. However, these analyses may be influenced by sexual dimorphism, since males and females are marked by significant biological differences that can modify the gene expression of economically important traits [12]. Therefore, researchers should take into account the conceptuses sex and sex ratios in their research before drawing any conclusion [13].

In mammals, sex determination is basically organized in four stages: 1) Determination of the chromosomal sex, which is established at fertilization when a Y- or a X-chromosome sperm fuses with the oocyte to determine the zygote genetic sex; 2) the differentiation of gonads into testicles or ovaries from the presence or absence of sex-determining region Y gene (SRY) at the critical time window during embryonic development; 3) the differentiation of male and female internal and external genitalia from the non-sexually dimorphic structures present

in the embryo, and 4) anatomical and physiological differences, i.e. sexual differentiation. Thus, all sex determination stages are mainly related to general gonadal function [14,15].

At early pig development, the absence of gonadal phenotype may impair the accurate visual sex determination, since the tunica albuginea, a marker of testis formation, is histologically identified only at 27 days [16] and the beginning of testicular descent occurs around 60 days [17]. Therefore, if research projects are designed at early development stages in which male and female conceptuses are not phenotypically divergent and there is response to sexual dimorphism, the sex determination could be performed using alternative tools rather than phenotypic differences.

Several studies have demonstrated how to identify the conceptuses sex at early stages of development using molecular techniques, as PCR (polymerase chain reaction) from Y chromosome-linked genes [13,18–20]. The PCR provides sensitive, precise, rapid, and reliable results [20]. However, if a RNA-seq study has already been performed on conceptuses to obtain a broader knowledge of the transcriptome under experimental conditions (e.g. inclusion of additives on maternal diets), these data can also be applied to determine the conceptuses sex with no need for extra techniques. A similar approach has been used by Petropoulos et al. [21] to determine the sex of human embryos through single-cell sequencing.

Therefore, considering that in early prenatal pig development male and female conceptuses are not phenotypically divergent and the sexual dimorphism

may affect the transcriptome analyses, we aimed to demonstrate that the sex of 25 and 35 days-old conceptuses can be determined by read counting of genes mapped on Y chromosome using already available RNA-seq data.

2. Materials and Methods

2.1. Experimental Animals and Design

The experimental protocols used in this study have followed ethical principles in animal research (CONCEA, 2016) and were approved by the Ethical Committee on Animal Use of the Universidade Federal de Viçosa (UFV), MG, Brazil [protocol # 06/2017].

All experimental protocols were performed according to the experimental design described in Costa et al. [22]. Briefly, 24 hours after the second insemination, 11 gilts received a basal diet for pregnant animals without supplementation (CONT) and 12 gilts received a CONT diet supplemented with 1.0% L-arginine (ARG), and two gestational ages were considered (25 and 35 days). From these gilts, 20 became pregnant. Additional details regarding reproductive management and nutritional information of the diets have been previously described in Costa et al. [22].

At 25 days of gestation, five females of CONT (n = 5) and five females of ARG (n = 5) were rendered unconscious using head-only electrical stunning (240V, 1.3A) and immediately exsanguinated. The same procedure was followed for four females of CONT (n = 4) and six females of ARG (n = 6) at 35 days of gestation. After slaughter, four conceptuses were collected per female at each gestational age, totaling in average 20 conceptuses per treatment.

The conceptuses were quickly washed with PBS (Phosphate Buffered Saline) solution, individually identified, stored in liquid nitrogen and transported to the Animal Biotechnology Laboratory (LABTEC) at the Department of Animal Science, UFV. At LABTEC, each conceptus (embryos from gilts slaughtered at 25 days and fetuses from gilts slaughtered at 35 days) was entirely and separately macerated in liquid nitrogen. Although an average of five gilts was initially used per treatment in the original trial described in Costa et al. [22], subsequent analyses were performed on a subset of gilts (n = 3) per treatment, since at least three biological replicates are recommended for describing results in RNA-seq experiments [9,23]. Moreover, three embryos from each CONT gilt (n = 3 gilts) and three embryos from each ARG gilt (n = 3 gilts) at 25 days of gestation (25DC and 25DA, respectively) and three fetuses from each CONT gilt (n = 3 gilts) and three fetuses from each ARG gilt (n = 3 gilts) at 35 days of gestation (35DC and 35DA, respectively), totaling 36 samples, were randomly chosen for RNA-seq analysis. Afterwards, these samples were transported in liquid nitrogen to the Animal Genetics Laboratory at the Embrapa Swine and Poultry National Research Center, Concordia, SC, Brazil, for further RNA extraction and library preparation.

2.2. RNA Extraction and Library Preparation

Total RNA extraction of the 25- and 35-day-old conceptuses from the ARG and CONT females was performed with TRIzol (Invitrogen, San Diego, CA, USA). The macerated conceptuses (100 mg) and TRIzol (1 mL) were mixed with vortex and then incubated for 5 minutes at room temperature (RT, 25 °C). Then, 200 µL of chloroform were added, shaking vigorously for 15 seconds and incubated at RT

for 5 minutes. Centrifugation was performed at 11,000× g at 4 °C for 15 minutes. Approximately 600 µL of the clear upper aqueous phase containing only RNA were carefully removed and transferred to a new tube, and 600 µL of 70% ethanol were added and homogenized by inversion. This volume was added to the silica column RNeasy mini kit (Qiagen, Hilden, Germany) and centrifuged for 15 seconds at 8000× g. The eluate was discarded and 700 µL RW1 buffer were added, followed by centrifugation for 15 seconds at 8000× g. Two washes with 500 µL RPE buffer were done and, finally, RNAs were eluted in 50 µL of RNase free water.

After RNA extraction, the quantification was performed in a QUBIT fluorimeter (Thermo Scientific, Waltham, MA, USA) and the integrity was determined in 1.0% agarose gel. In addition, the Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) was used for integrity measurement, in which samples with RNA integrity number (RIN) higher than eight were used for library preparation. In this sense, one fetus sample was excluded from subsequent analysis due to low RIN score (RIN < 8). Therefore, the 35 remaining conceptuses samples (nine ARG and nine CONT at 25 days and eight ARG and nine CONT at 35 days) were submitted to RNA-seq library preparation using the TruSeq Stranded mRNA Library Prep Kit (Illumina, Inc., San Diego, CA, USA), followed by purification of the poly-A tail using 2 µg of total RNA, according to the manufacturer´s recommendations.

2.3. Sequencing, Quality Control and Mapping

The libraries were sequenced in Illumina HiSeq2500 (Illumina, Inc.; San Diego, CA, USA), following the 2 × 100 bp paired-end protocol, at the Functional

Genomics Center, ESALQ, Universidade de São Paulo, Piracicaba, SP, Brazil. The FASTQ files were deposited in the SRA database, with Bioproject number PRJNA576701 and Biosample numbers SAMN13003023, SAMN13003024, SAMN13003025, SAMN13003026, SAMN13003027, SAMN13003028, SAMN13003029, SAMN13003030, SAMN13003031, SAMN13003032, SAMN13003033, SAMN13003034, SAMN13003035, SAMN13003036, SAMN13003037, SAMN13003038, SAMN13003039, SAMN13003040, SAMN13003041, SAMN13003042, SAMN13003043, SAMN13003044, SAMN13003045, SAMN13003046, SAMN13003047, SAMN13003048, SAMN13003049, SAMN13003050, SAMN13003051, SAMN13003052, SAMN13003053, SAMN13003054, SAMN13003055, SAMN13003056 and SAMN13003057.

The BAQCOM pipeline [24] was used to perform quality control (QC), mapping and reads counting. BAQCOM uses a set of software to analyze RNA-seq data, as Trimmomatic [25], version 0.38, to identify and remove adapter and low quality sequences, HISAT2 [26], version 0.11.2, to map reads against the genome and HTseq-count [27], version 0.11.2, to count reads in features. Using BAQCOM pipeline, only reads with Phred quality ≥ 20 and length ≥ 70 pb were mapped against the pig reference genome (Sus scrofa, v. 11.1) and the reads counting for each feature was based on Ensembl annotation release 95.

2.4. Sex Identification

The conceptuses sex determination (n = 18 embryos and n = 17 fetuses) was performed using discrepant reads counts per million (CPM) of Y chromosome-

linked genes from the conceptuses RNA-seq dataset. After the samples classification as males or females, the HTseq files containing all samples transcripts were renamed and used to build a multi-dimensional scale (MDS) plot using plotMDS function of EdgeR package [28] of R software [29] to evaluate conceptuses dispersion.

3. Results

Averages of 15.3 and 14.4 million reads/samples were generated for embryos and fetuses, respectively. After the data QC, averages of 13.6 million reads/embryos and 12.9 million reads/fetuses remained for further analyses. More than 98% of the reads were mapped against the pig reference genome (*Sus scrofa*, v.11.1), with 83.8% and 79.8% of the reads counted into genes for embryos and fetuses, respectively. An average of 0.05% of the reads was uniquely mapped on Y chromosome. At 25 and 35 days of age, male and female conceptuses could not be phenotypically differentiated, since gonadal phenotype could not be observed (Figure S1).

From 79 Y chromosome-linked genes, ten showed discrepant reads CPM among samples, i.e, high and low reads CPM values (Tables 1 and 2), and were selected for sex determination: *DDX3Y (*ATP-dependent RNA helicase DDX3X), *KDM5D* (Lysine demethylase 5D*), ZFY (Z*inc Finger Protein, Y-Linked), *EIF2S3Y* (eukaryotic translation initiation factor 2 subunit 3), *EIF1AY (*eukaryotic translation initiation factor 1A, Y-linked), *LOC110255320* (lysine-specific demethylase 6A-like), *LOC110257894* (gamma-taxilin-like), *LOC396706* (U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2),

LOC100625207 (probable ubiquitin carboxyl-terminal hydrolase FAF-X) and *LOC110255257* (oral-facial-digital syndrome 1 protein-like). The CPM sum for these genes (ΣCPM_{chrY}) was calculated for each conceptus in order to determine its sex. In this way, conceptuses that presented ΣCPM_{chrY} greater than 400 were classified as males and those with ΣCPM_{chrY} below 2 were classified as females. Male conceptuses showed mean (standard deviation), minimum and maximum ΣCPM_{chrY} of 499.08 (44.74), 427.29, and 573.09, respectively (Table 1), while females showed mean (standard deviation), minimum and maximum ΣCPM_{chrY} of 0.38 (0.42), 0.00 and 1.75, respectively (Table 2). In the MDS plot, a segregation of conceptuses into two distinct sex groups was observed (Figure 1).

Table 1. Reads counts per million of 10 genes located in the Y chromosome of male pig conceptuses.

Conceptuses Sample_ID [1]	Counts Per Million (CPM)										
	DDX3Y [2]	*KDM5D*	*ZFY*	*EIF2S3Y*	*EIF1AY*	*LOC110255320*	*LOC110257894*	*LOC396706*	*LOC100625207*	*LOC110255257*	ΣCPM_{chrY} [3]
E1-25DA-M	151.46	7.52	24.81	162.23	40.64	12.04	9.71	1.53	61.73	15.17	486.84
E2-25DA-M	166.54	7.64	31.68	197.73	45.16	12.30	11.81	1.95	61.00	17.72	553.53
E3-25DC-M	167.21	9.56	28.67	211.97	45.86	13.23	12.79	2.21	66.52	15.07	573.09
E4-25DC-M	170.76	10.15	29.78	189.79	46.39	11.43	12.41	2.11	71.81	17.90	562.53
E5-25DC-M	149.28	6.56	24.94	180.78	36.55	13.47	10.47	1.43	55.29	16.46	495.23
E6-25DC-M	156.64	9.19	30.92	192.23	40.76	12.14	10.99	2.30	68.56	16.24	539.97
E7-25DC-M	166.10	9.43	33.57	192.23	44.92	12.31	11.43	1.84	71.78	17.03	560.64
F1-35DC-M	193.41	5.35	26.07	130.67	44.65	9.79	6.25	1.23	52.46	13.16	483.04
F2-35DC-M	167.28	7.11	28.43	135.38	39.84	10.25	7.19	1.41	56.20	15.04	468.13
F3-35DC-M	168.09	7.05	31.32	141.53	37.80	10.96	7.26	1.78	58.09	13.03	476.91
F4-35DA-M	157.51	8.17	26.77	119.45	33.45	9.30	6.46	2.06	51.27	12.85	427.29
F5-35DA-M	200.33	8.04	24.02	126.13	45.09	12.20	6.56	2.13	45.83	12.20	482.53
F6-35DA-M	187.01	6.47	24.02	128.61	41.58	10.35	8.01	2.02	45.54	13.43	467.04
F7-35DA-M	180.78	7.06	29.27	133.42	41.14	11.74	6.57	2.10	50.78	13.20	476.06
F8-35DC-M	178.15	10.68	33.04	155.57	43.49	11.29	9.23	2.06	62.94	17.32	523.77
F9-35DA-M	143.33	10.28	27.83	135.33	31.42	12.33	8.65	1.55	53.79	12.73	437.24
F10-35DA-M	173.51	10.40	26.69	128.32	36.71	12.62	6.88	1.91	58.19	15.37	470.60

[1]Conceptuses Sample Identification: Seven male embryos samples from supplemented (E-25DA-M) and non-supplemented (E-25DC-M) gilts; ten male fetuses samples from supplemented (F-35DA-M) and non-supplemented (F-35DC-M) gilts; [2]*DDX3Y*: ATP-dependent RNA

helicase DDX3X; *KDM5D*: Lysine demethylase 5D; *ZFY*: Zinc Finger Protein Y-Linked; *EIF2S3Y*: eukaryotic translation initiation factor 2 subunit 3; *EIF1AY*: eukaryotic translation initiation factor 1A, Y-linked; *LOC110255320*: Lysine-specific demethylase 6A-like; *LOC110257894*: Gamma-taxilin-like; *LOC396706*: U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2; *LOC100625207*: Probable ubiquitin carboxyl-terminal hydrolase FAF-X; *LOC110255257*: Oral-facial-digital syndrome 1 protein-like; [3]ΣCPM_{chrY}: sum of reads CPM of 10 Y chromosome-linked genes in male conceptuses.

Table 2. Reads counts per million of 10 Y chromosome-linked genes that showed reads counting in female pig conceptuses.

Conceptuses Sample_ID [1]	Counts Per Million (CPM)										
	DDX3Y [2]	*KDM5D*	*ZFY*	*EIF2S3Y*	*EIF1AY*	*LOC110255320*	*LOC110257894*	*LOC396706*	*LOC100625207*	*LOC110255257*	ΣCPM_{chrY} [3]
E1-25DA-F	0.00	0.07	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.14
E2-25DA-F	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.07	0.14
E3-25DC-F	0.00	0.00	0.07	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.14
E4-25DC-F	0.28	0.00	0.07	0.07	0.07	0.00	0.00	0.00	0.00	0.00	0.49
E5-25DC-F	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.07	0.15	0.37
E6-25DC-F	0.00	0.00	0.14	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.21
E7-25DC-F	0.00	0.00	0.15	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.23
E8-25DC-F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E9-25DC-F	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.07	0.14
E10-25DC-F	0.00	0.00	0.08	0.08	0.08	0.00	0.00	0.00	0.00	0.00	0.24
E11-25DA-F	0.13	0.00	0.07	0.13	0.07	0.00	0.00	0.00	0.20	0.00	0.60
F1-35DA-F	0.08	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.24
F2-35DA-F	0.00	0.08	0.23	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.39

F3-35DA-F	0.16	0.00	0.00	0.25	0.08	0.00	0.00	0.00	0.08	0.00	0.57
F4-35DC-F	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07
F5-35DA-F	0.71	0.00	0.32	0.48	0.08	0.00	0.00	0.00	0.16	0.00	1.75
F6-35DA-F	0.22	0.07	0.14	0.29	0.07	0.00	0.00	0.00	0.22	0.00	1.01
F7-35DA-F	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07

[1]Conceptuses Sample Identification: Eleven female embryos samples from supplemented (E-25DA-F) and non-supplemented (E-25DC-F) gilts; seven female fetuses samples from supplemented (F-35DA-F) and non-supplemented (F-35DC-F) gilts; [2]*DDX3Y*: ATP-dependent RNA helicase DDX3X; *KDM5D*: Lysine-specific Demethylase 5D; *ZFY*: Zinc Finger Protein Y-Linked; *EIF2S3Y*: eukaryotic translation initiation factor 2 subunit 3; *EIF1AY*: eukaryotic translation initiation factor 1A, Y-linked; *LOC110255320*: Lysine-specific demethylase 6A-like; *LOC110257894*: Gamma-taxilin-like; *LOC396706*: U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit-related protein 2; *LOC100625207*: Probable ubiquitin carboxyl-terminal hydrolase FAF-X; *LOC110255257*: Oral-facial-digital syndrome 1 protein-like; [3]ΣCPM_{chrY}: sum of reads CPM of 10 Y chromosome-linked genes in females conceptuses.

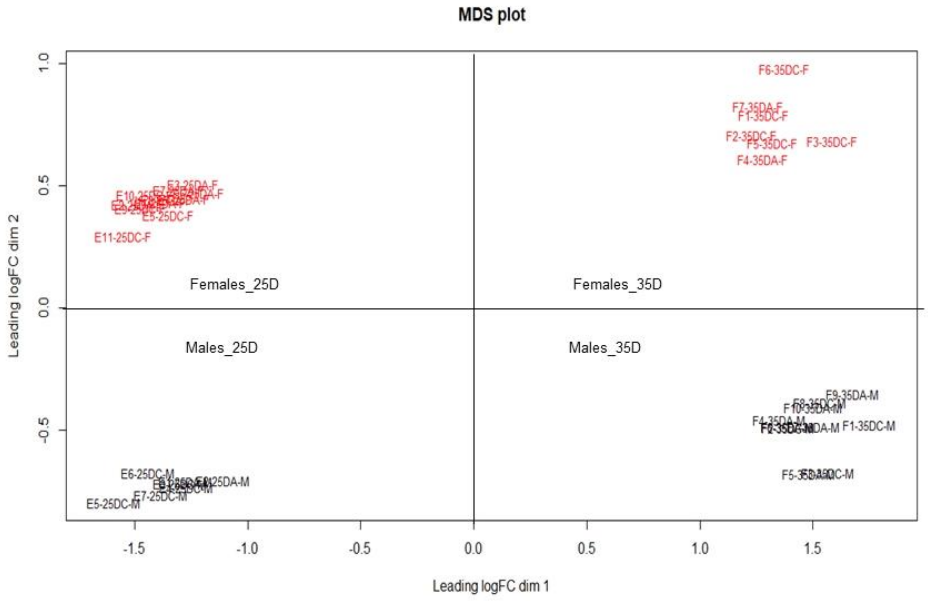


Figure 1. Multi-dimensional scale (MDS) plot showing segregation of pig conceptuses (25D: embryos at 25 days-old and 35D: fetuses at 35 days-old) into two distinct sex groups. Red: Female conceptuses; Black: Male conceptuses; Samples identification: male embryos from supplemented (E-25DA-M) or non-supplemented (E-25DC-M) gilts; male fetuses from supplemented (F-35DA-M) or non-supplemented (F-35DC-M) gilts; female embryos from supplemented (E-25DA-F) or non-supplemented (E-25DC-F) gilts; female fetuses from supplemented (F-35DA-F) or non-supplemented (F-35DC-F) gilts.

4. Discussion

The control of animals sex ratio is desirable in livestock production [20], since the expression of genes that affect a wide range of economically important traits may be influenced by sexual dimorphism [12]. Considering that the sex determination is equivalent to testis determination [30], the testis could be easily used to determine male and female individual rates. However, in pig conceptuses, the tunica albuginea, a marker of testis formation, is histologically identified only at 27 days [16] and the beginning of testicular descent occurs around 60 days [17]. In this context, we demonstrated that it is possible to determine the sex of pig embryos (25 days) and fetuses (35 days) using reads CPM of ten Y chromosome-

linked genes through RNA-seq analysis, with no need to use other molecular tools for this purpose, such as PCR. Although most often considered a cheaper and faster technique than RNA-seq, PCR would represent another cost source in this study, since RNA-seq analyses had been previously performed aiming to test a hypothesis on a maternal nutrition experiment [22].

Several studies have also used Y chromosome-linked genes to determine the sex of pigs at early development stages [13,18,19], however, all of them have applied DNA-based approaches. The Y chromosome harbors genes that are essential for testis development and function, such as the master gene for testis determination (*SRY)*, which is expressed during a critical period of embryonic development, and other genes that are important for spermatogenic function, as the *ZFY*, *EIF1AY* and *KDM5D*. The importance of these Y-linked genes during male development has been elucidated for mammals [31–34]. These genes are located in the non-recombining region of the Y chromosome (NRY), which is also known as the male-specific Y (MSY) region, transmitted from fathers to sons without recombination with the X chromosome [35]. The MSY region contains only 27 genes coding for distinct proteins, as the *SRY*, *EIF1AY*, *KDM5D*, *ZFY*, *EIF2S3Y,* and *DDX3Y* [31,35,36].

In this study, among all abovementioned genes, *SRY* was not used to differentiate male and female conceptuses (data not shown). This gene seems to follow a strict expression pattern during mammalian embryonic development. In mouse, its expression starts at 10.5 days post coitum (dpc), reaches a peak at 11.5 dpc and then wanes by 12.5 dpc [32,37]. In pigs, the *SRY* expression is detectable

in the embryos genital ridge around 21 and 23 days, with faint expression at 26 days and absence of expression at 31 days [38]. Consequently, both 25 and 35 days of pig prenatal periods do not comprise the main *SRY* expression phase, reinforcing the critical time window during embryonic development in which *SRY* is expressed.

Subsequently to the sex determination by the *SRY* gene, other male-specific genes are important to maintain testicular development and spermatogenesis [39]. According to Bellott et al. [40], during spermatogenesis, each stage of the molecular central dogma is regulated by the following Y-linked genes: *KDM5D*, *ZFY,* and *EIF1AY*. The *KDM5D* gene encodes a histone demethylase enzyme that acts removing tri- and di-methylations of lysine 4 of histone H3 at the start site of transcription in actively transcribed genes [41]. The transcription factor *ZFY*, the first coding gene identified in the human Y chromosome, regulates the transcription of a number of Y-linked genes [31] and mediates multiple aspects of spermatogenesis and reproduction, such as morphology, motility, capacitation, acrosome reaction, and oocyte activation, as well as chromosomal aberrations [42]. The translation initiation factor *EIF1AY* is requested for a high rate of protein biosynthesis and the absence of this gene may contribute to failures in the spermatogenic process [43].

Another important translation initiation factor during spermatogenesis is the *EIF2S3Y* gene, located in the short arm of the Y chromosome. This gene has been identified as a key mouse-specific regulator of spermatogonia proliferation and differentiation [44]. Finally, the *DDX3Y* gene, along with *KDM5D* and *EIF1AY*, are

azoospermia factors (AZFa) [45]. Azoospermia is defined as the absence of sperm in at least two different ejaculates [31]. The *DDX3Y* is the major AZFa gene and belongs to a highly conserved subfamily of the DEAD-box RNA helicase family [33]. This gene family is involved in oogenesis and spermatogenesis, silencing mobile elements and other repetitive genomic regions in germinal tissues, and is important for gonads formation during embryo development [46]. Therefore, the *KDM5D*, *ZFY*, *EIF1AY*, *EIF2S3Y,* and *DDX3Y* genes showed discrepant reads CPM among samples and have important spermatogenic and gonadal functions, being crucial to differentiate male and female pig conceptuses in the present study. Similar pattern has already been observed in humans blastocyst [21], which highlights the importance of using pigs as an animal model to study sex determination throughout human prenatal development.

The reads CPM observed for the Y-linked genes in female conceptuses may be considered technical artifacts from RNA-seq, since mammalian X and Y chromosomes share regions of high sequence similarity, causing mismapping of short reads to a reference genome [47]. In this context, Ballouz et al. [48] found reads of female samples mapped in Y chromosome, which was also characterized as reads mismapping. Although considered technical artifacts, these Y chromosome reads CPM found in females were not a major issue in the present study, since the discrepant reads CPM among samples for the ten Y chromosome-linked genes were sufficient to differentiate males and females.

The *LOC110255257*, *LOC100625207*, *LOC110255320*, *LOC110257894,* and *LOC396706* genes are located in the Y chromosome, however, they have not

yet been fully characterized in *Sus scrofa* genome. According to Skinner et al. [49], there are few data available on the porcine Y chromosome. Therefore, although these genes are located on this chromosome and have shown contribution to sex determination of pig conceptuses due to the discrepant reads CPM among samples, their functions will not be addressed in this study. Further studies might be performed to characterize these Y-linked genes in pigs.

5. Conclusions

We have demonstrated that sex identification can be performed from read counts per million of genes mapped in the Y chromosome at early stages of pig development using already available conceptuses RNA-seq data, with no need for extra techniques. Additionally, the results reinforce that molecular sex determination is a mechanism conserved across mammalian species, since the role of these genes is also important during male development in humans, highlighting the applicability of pigs as an animal model to study sex determination during human prenatal development.

Supplementary Materials: The following is available online at http://www.mdpi.com/2073-4425/10/12/1010/s1, Figure S1: 25 days-old pig embryo (left) and 35 days-old pig fetus (right) do not present evident gonadal phenotype.

Author Contributions: S.A.T., A.M.G.I., M.C.L, J.O.P., and S.E.F.G.; methodology, S.A.T., A.M. G.I., M.E.C., H.C.O., L.L.C.; software, S.A.T., A.M.G.I., M.E.C., H.C.O.; formal Analysis, S.A.T., A.M.G., M.E.C., H.C.O. and D.B.D.M.; writing—Original draft preparation S.A.T.; writing—Review and editing, A.M.G.I., M.C.L, J.O.P., M.E.C., H.C.O., D.B.D.M., K.A.C. and S.E.F.G; supervision,

S.E.F.G.; project administration, S.E.F.G.; funding acquisition, M.C.L., J.O.P., S.E.F.G.

Funding: This research was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Grant # 423558/2018-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Grant numbers PROEX 0577/2018 and 320.020.170.11pq), Instituto Nacional de Ciência e Tecnologia-Ciência Animal (INCT-CA, Grant # 465377/2014-9) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Grant # CVZ APQ 03434/16).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Godfrey, K.M.; Barker, D.J.P. Fetal programming and adult health. *Public Health Nutr.* **2001**, *4*, 611–624.
2. Li, J.; Xia, H.; Yao, W.; Wang, T.; Li, J.; Piao, X.; Thacker, P.; Wu, G.; Wang, F. Effects of arginine supplementation during early gestation (day 1 to 30) on litter size and plasma metabolites in gilts and sows 1. *J. Anim. Sci.* **2015**, *93*, 5291–5303.
3. Wu, G.; Bazer, F.W.; Cudd, T.A.; Meininger, C.J.; Spencer, T.E. Recent Advances in Nutritional Sciences Maternal Nutrition and Fetal Development. *J. Nutr.* **2004**, *134*, 2169–2172.
4. Zhang, S.; Heng, J.; Song, H.; Zhang, Y.; Lin, X.; Tian, M.; Chen, F.; Guan, W. Role of Maternal Dietary Protein and Amino Acids on Fetal Programming, Early Neonatal Development, and Lactation in Swine. *Animals* **2019**, *9*, 19.
5. Assis, K.; Botelho, D.; Marques, D.; Filardi, C.; Campos, D.; Saraiva, A.; Domingos, J.; Eliza, S.; Guimarães, F. Nutrition in fluence on sow reproductive performance and conceptuses development and survival: A review about L-arginine supplementation. *Livest. Sci.* **2019**, *228*, 97–103.
6. Peñagaricano, F.; Wang, X.; Rosa, G.J.M.; Radunz, A.E.; Khatib, H. Maternal nutrition induces gene expression changes in fetal muscle and adipose tissues in sheep. *BMC Genom.* **2014**, *15*, 1034.
7. Verma, N.; Rettenmeier, A.W.; Schmitz-Spanke, S. Recent advances in the use of Sus scrofa (pig) as a model system for proteomic studies. *Proteomics* **2011**, *11*, 776–793.
8. Papatheodorou, I.; Fonseca, N.A.; Keays, M.; Tang, Y.A.; Barrera, E.; Bazant, W.; Burke, M.; Füllgrabe, A.; Fuentes, A.M.P.; George, N.; et al. Expression Atlas: Gene and protein expression across multiple studies and organisms. *Nucleic Acids Res.* **2018**, *46*, D246–D251.

9. Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-cabrero, D.; Cervera, A.; Mcpherson, A.; Szcze, W.; Gaffney, D.J.; Elo, L.L.; Zhang, X. A survey of best practices for RNA-seq data analysis. *Genome Biol.* **2016**, *17*, 13.
10. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63.
11. Sanglard, L.P.; Nascimento, M.; Moriel, P.; Sommer, J.; Ashwell, M.; Poore, M.H.; Duarte, M.D.S.; Serão, N.V.L. Impact of energy restriction during late gestation on the muscle and blood transcriptome of beef calves after preconditioning. *BMC Genom.* **2018**, *19*, 702.
12. Cui, L.; Zhang, J.; Ma, J.; Guo, Y.; Li, L.; Xiao, S.; Ren, J.; Yang, B.; Huang, L. Sexually dimorphic genetic architecture of complex traits in a large-scale F_2 cross in pigs. *Genet. Sel. Evol.* **2014**, *46*, 76.
13. Blanes, M.S.; Tsoi, S.C.M.; Dyck, M.K. Accurate and Phenol Free DNA Sexing of Day 30 Porcine Embryos by PCR. *J. Vis. Exp.* **2016**, *108*, e53301.
14. Miura, K.; Tomita, A.; Kanai, Y. Sex Determination and Differentiation in Mammals. In *Reproductive and Developmental Strategies*; Kazuya, K., Takeshi, K., Yasuhiro, I., Mariko, K., Eds.; Springer: Tokyo, Japan, 2018; pp. 407–433, ISBN 9784431566090.
15. Goodfellow, P.N.; Darling, S.M. Genetics of sex determination in man and mouse. *Development* **1988**, *102*, 251–258.
16. Pelliniemi, L.J. Ultrastructure of the early ovary and testis in pig embryos. *Am. J. Anat.* **1975**, *144*, 89–111.
17. Hyttel, P.; Sinowatz, F.; Vejlsted, M.; Betteridge, K. *Essentials of Domestic Animal Embryology*; Edwards, R., Rodenhuis, J., Betteridge, K., Eds.; Saunders: *Philadelphia*, PA, USA, 2010; ISBN 978-0-7020-2899-1.
18. Tsoi, S.; Blanes, M.; Chen, T.Y.; Langendijk, P.; Athorn, R.; Foxcroft, G.; Dyck, M. Identification of differentially expressed genes in sexed pig embryos during post-hatching development in primiparous sows exposed to differing intermittent suckling and breeding strategies. *Genom. Data* **2016**, *9*, 30–34.
19. Bick, J.T.; Flöter, V.L.; Robinson, M.D.; Bauersachs, S.; Ulbrich, S.E. Small RNA-seq analysis of single porcine blastocysts revealed that maternal estradiol-17beta exposure does not affect miRNA isoform (isomiR) expression. *BMC Genom.* **2018**, *19*, 590.
20. Han, S.H.; Yang, B.C.; Ko, M.S.; Oh, H.S.; Lee, S.S. Length difference between equine ZFX and ZFY genes and its application for molecular sex determination. *J. Assist. Reprod. Genet.* **2010**, *27*, 725–728.
21. Petropoulos, S.; Edsgärd, D.; Reinius, B.; Deng, Q.; Panula, S.P.; Codeluppi, S.; Reyes, A.P.; Linnarsson, S.; Sandberg, R.; Lanner, F. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **2016**, *165*, 1012–1026.
22. Costa, K.A.; Saraiva, A.; Guimaraes, J.D.; Marques, D.B.D.; Machado-neves, M.; Barbosa, L.M.R.; Villadiego, F.A.C.; Veroneze, R.; de Oliveira, L.F.; Garcia, I.S.; et al. Dietary L-arginine supplementation during early gestation of gilts affects conceptuses development. *Theriogenology* **2019**, *140*, 62–71.

23. Nie, H.; Wang, Y.; Su, Y.; Hua, J. Exploration of miRNAs and target genes of cytoplasmic male sterility line in cotton during flower bud development. *Funct. Integr. Genom.* **2018**, *18*, 457–476.
24. BAQCOM: Bioinformatics Analysis for Quality Control and Mapping. Available online: https://github.com/hanielcedraz/BAQCOM (https://github.com/hanielcedraz/BAQCOM accessed on April, 15th, 2019).
25. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120.
26. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360.
27. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169.
28. Robinson, M.D.; Mccarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140.
29. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
30. Koopman, P.; Gubbay, J.; Vivian, N.; Goodfellow, P.; Lovell-Badge, R. Male development of chromosomally female mice transgenic for Sry. *Nature* **1991**, *351*, 117–121.
31. Krausz, C.; Casamonti, E. Spermatogenic failure and the Y chromosome. *Hum. Genet.* **2017**, *136*, 637–655.
32. Kashimada, K.; Koopman, P. Sry: The master switch in mammalian sex determination. *Development* **2010**, *137*, 3921–3930.
33. Rauschendorf, M.A.; Zimmer, J.; Ohnmacht, C.; Vogt, P.H. *DDX3X*, the X homologue of AZFa gene *DDX3Y*, expresses a complex pattern of transcript variants only in the male germ line. *Mol. Hum. Reprod.* **2014**, *20*, 1208–1222.
34. Peneder, P.; Wallner, B.; Vogl, C. Exchange of genetic information between therian X and Y chromosome gametologs in old evolutionary strata. *Ecol. Evol.* **2017**, *7*, 8478–8487.
35. Fan, H.; Hu, Y.; Shan, L.; Yu, L.; Wang, B.; Li, M.; Wu, Q.; Wei, F. Synteny search identifies Carnivore Y chromosome for evolution of male specific genes. *Integr. Zool.* **2018**, *14*, 224–234.
36. Maan, A.A.; Eales, J.; Akbarov, A.; Rowland, J.; Xu, X.; Jobling, M.A.; Charchar, F.J.; Tomaszewski, M. The y chromosome: A blueprint for men's health? *Eur. J. Hum. Genet.* **2017**, *25*, 1181–1188.
37. Hacker, A.; Capel, B.; Goodfellow, P.; Lovell-Badge, R. Expression of Sry, the mouse sex determining gene. *Development* **1995**, *121*, 1603–1614.
38. Daneau, I.; Ethier, J.-F.; Lussier, J.G.; Silversides, D.W. Porcine SRY Gene Locus and Genital Ridge Expression. *Biol. Reprod.* **2005**, *55*, 47–53.
39. Akimoto, C.; Kitagawa, H.; Matsumoto, T.; Kato, S. Spermatogenesis-specific association of SMCY and MSH5. *Genes Cells* **2008**, *13*, 623–633.
40. Bellott, D.W.; Hughes, J.F.; Skaletsky, H.; Brown, L.G.; Pyntikova, T.; Cho, T.; Koutseva, N.; Zaghlul, S.; Graves, T.; Rock, S.; et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **2014**, *508*, 494–499.

41. Rasmussen, P.B.; Staller, P. The KDM5 family of histone demethylases as targets in oncology drug discovery. *Epigenomics* **2014**, *6*, 277–286.
42. Nakasuji, T.; Ogonuki, N.; Chiba, T.; Kato, T.; Shiozawa, K.; Yamatoya, K.; Tanaka, H.; Kondo, T.; Miyado, K.; Miyasaka, N.; et al. Complementary Critical Functions of *Zfy1* and *Zfy2* in Mouse Spermatogenesis and Reproduction. *PLoS Genet.* **2017**, *13*, e1006578.
43. Kleiman, S.E.; Yogev, L.; Hauser, R.; Botchan, A.; Maymon, B.B.S.; Paz, G.; Yavetz, H. Expression profile of AZF genes in testicular biopsies of azoospermic men. *Hum. Reprod.* **2007**, *22*, 151–158.
44. Mazeyrat, S.; Saut, N.; Grigoriev, V.; Mahadevaiah, S.K.; Ojarikre, O.A.; Rattigan, Á.; Bishop, C.; Eicher, E.M.; Mitchell, M.J.; Burgoyne, P.S. A Y-encoded subunit of the translation initiation factor Eif2 is essential for mouse spermatogenesis. *Nat. Genet.* **2002**, *29*, 49–53.
45. Mitchell, M.J.; Metzler-Guillemain, C.; Toure, A.; Coutton, C.; Arnoult, C.; Ray, P.F. Single gene defects leading to sperm quantitative anomalies. *Clin. Genet.* **2017**, *91*, 208–216.
46. Kotov, A.A.; Akulenko, N.V.; Kibanov, M.V.; Olenina, L.V. DEAD-Box RNA Helicases in Animal Gametogenesis. *Mol. Biol.* **2014**, *48*, 16–28.
47. Webster, T.H.; Couse, M.; Grande, B.M.; Karlins, E.; Phung, T.N.; Richmond, P.A.; Whitford, W.; Wilson, M.A. Identifying, understanding, and correcting technical artifacts on the sex chromosomes in next-generation sequencing data. *GigaScience* **2019**, *8*, giz074.
48. Ballouz, S.; Dobin, A.; Gingeras, T.R.; Gillis, J. The fractured landscape of RNA-seq alignment: The default in our STARs. *Nucleic Acids Res.* **2018**, *46*, 5125–5138.
49. Skinner, B.M.; Sargent, C.A.; Churcher, C.; Hunt, T.; Herrero, J.; Loveland, J.E.; Dunn, M.; Louzada, S.; Fu, B.; Chow, W.; et al. The pig X and Y chromosomes: Structure, sequence and evolution. *Genome Res.* **2016**, *26*, 130–139.

CHAPTER 3

RNA-sequencing reveals differences in expression of energy metabolism genes in embryos from dietary L-arginine supplemented gilts*

Susana A. Teixeira[a], Adriana M. G. Ibelli[b], Maurício E. Cantão[b], Haniel C. de Oliveira[a], Daniele B. D. Marques[a], Mônica C. Ledur[b], Jane de O. Peixoto[b], Karine A. Costa[a], Luiz Lehmann Coutinho[c], José Domingos Guimarães[d], Alysson Saraiva[a], Simone E. F. Guimarães[a]**

[a]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil;

[b]Animal Genetics Laboratory, Embrapa Swine and Poultry Nacional Research Center, Concordia, SC, 89715-899, Brazil;

[c]Functional Genomics Center, ESALQ, Universidade de São Paulo, Piracicaba, SP, 89715-899, Brazil;

[d]Department of Veterinary Medicine, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil;

[**]Corresponding author. Email: sfacioni@ufv.br Tel.: +55 31 36124625

---

*This article was published as a part of research article in Molecular Reproduction & Development
Received: 27 March 2020
Accepted: 12 June 2020
DOI: 10.1002/mrd.23397

ABSTRACT: Dietary L-arginine supplementation has been studied as a nutritional strategy to improve reproductive performance of pregnant sows, since arginine is a conditionally essential amino acid. However, reports addressing the molecular mechanisms that mediate the supplementation effects on fetal programming are still scarce. We hypothesized that maternal dietary L-arginine supplementation changes the conceptuses' transcriptome. Therefore, we evaluated the molecular effects of dietary 1.0% L-arginine supplementation (ARG) of commercial pregnant gilts on global transcriptome profile of 25 and 35 days-old conceptuses (18 embryos and 17 fetuses, respectively) using RNA-sequencing analysis. The mitochondrial genes *ND1*, *ND2*, *CYTB*, *COX2*, *ATP8* and a *rRNA type* similar to 16*S rRNA* were down-regulated in ARG embryos, while *CYP1A1* gene was up-regulated (adjusted *P*-value≤0.05), indicating energy metabolism adaptation against greater nutrients supply mediated by oxide nitric effects, since down-regulation of mitochondrial genes is associated with lower ATP synthesis and up-regulation of detoxification gene suggests increase of toxic metabolites in embryonic environment. On the other hand, the 35 days old fetuses' transcriptome was not affected by L-arginine, providing support for further discussions regarding the optimal period of supplementation, in order to understand its impacts on swine prenatal life.

Keywords: embryos metabolism, fetal programming, maternal nutrition, nutrigenomics, prenatal development.

1. INTRODUCTION

The genetic selection for litter size has increased the proportion of IUGR (intrauterine growth restriction) piglets (J. Wang et al., 2017). The intrauterine environment, which provides molecules for placenta development, as well as for survival, growth and development of conceptuses, is related with IUGR and it is mainly determined by nutrients such as amino acids from maternal diet (Wu, Bazer, Cudd, Meininger, & Spencer, 2004; Zhang et al., 2019). Among these amino acids, arginine plays a key role in placental development and, consequently, in conceptuses growth (Wu, 2014; Wu et al., 2009, 2004; Wu, Bazer, Johnson, & Hou, 2018; Wu, Ott, Knabe, & Bazer, 1998).

Arginine is one of the most versatile amino acids in animal cells, acting on synthesis of several biologically active molecules, such as nitric oxide (NO), ornithine, polyamines (putrescine, spermidine, and spermine), creatine, and agmatine, besides its role on protein synthesis (Wu & Morris, 1998). Among arginine metabolites, NO and polyamines stimulate cell proliferation and migration, cellular remodeling, angiogenesis (the growth of new vessels from the existing vasculature), dilation of blood vessels and immune response (Costa et al., 2019; Keshet & Erez, 2018; Wu et al., 2009). During gestation, these metabolites play important roles on the control of placental development, vascularization and blood flow from the mother to the embryo, which may increase the supply of nutrients and oxygen to conceptuses (Bird, Zhang, & Magness, 2003; Wu et al., 2008). Due to these versatile functions, the arginine endogenous synthesis does not meet the cellular requirements under certain conditions, as pregnancy (Wu et al., 2018; Wu & Morris, 1998) and, therefore, it should be provided from diets for pregnant

mammals to support prenatal survival and growth (Li et al., 2010, 2014; Wu et al., 2009).

Maternal nutrition may affect progeny anatomy, physiology and metabolism during critical periods of embryonic and fetal developments (fetal programming) (Godfrey & Barker, 2001; Kwon & Kim, 2017) through regulation of cell signaling and epigenetic mechanisms (Ji et al., 2017), influencing the nutrient transport, angiogenesis and relaxation of vascular tissues, proliferation, migration and cell remodeling (Costa et al., 2019; Bazer, Spencer, Johnson, Burghardt, & Wu, 2009). In this context, several reports have shown the beneficial effects of maternal L-arginine supplementation on female reproductive performance (Gao et al., 2012; Mateo et al., 2007; Palencia et al., 2017; Wu et al., 2018). However, it is still unclear the molecular mechanisms that mediate L-arginine supplementation on fetal programming and transcriptome analysis by RNA-sequencing (RNA-seq) can be a powerful tool to measure molecular changes during development under different conditions (Z. Wang, Gerstein, & Snyder, 2009).

Based on maternal nutrition effects on fetal programming, we hypothesized that dietary L-arginine supplementation of gilts during gestation impacts the molecular mechanisms involved in conceptuses development. Therefore, we aimed to evaluate the effects of 1.0% dietary L-arginine supplementation of commercial pregnant gilts on global transcriptome profile of 25 and 35 days-old conceptuses using RNA-seq analysis.

2. MATERIAL AND METHODS

2.1 Animals and samples collection

The experimental protocol has followed ethical principles in animal research (CONCEA, 2016) and was approved by the Ethical Committee on Animal Use of *Universidade Federal de Viçosa* (UFV), MG, Brazil [protocol # 06/2017].

All protocols were performed according to the experimental design and reproductive management previously described in Costa et al. (2019). Briefly, after identification of the fourth estrus, 23 commercial gilts were inseminated 12 and 24 hours after the beginning of estrus. The first insemination day was considered day zero of gestation and the supply of experimental diets occurred 24 hours after the second insemination. A total of 11 gilts were fed a control diet (CON) and 12 gilts were fed the CON diet supplemented with 1.0% L-arginine (ARG), and two gestational ages were considered (25 and 35 days of gestation). Additional details regarding reproductive management and nutritional information have been described by Costa et al. (2019).

From the 23 inseminated gilts, 20 became pregnant. At 25 days of gestation, five females of CON and five females of ARG were rendered unconscious using head-only electrical stunning (240V, 1.3A) and immediately exsanguinated. The same procedure was followed for four CON gilts and six ARG gilts at 35 days of gestation.

Analyses from conceptuses (embryos from gilts slaughtered at 25 days and fetuses from gilts slaughtered at 35 days) samples were performed within each gestational age, considering a completely randomized design with two diets (CON and ARG) and five replicates on average per diet. After slaughter, four conceptuses were collected per female at each gestational age, totaling in average

20 conceptuses per diet. The conceptuses were quickly washed with PBS (Phosphate Buffered Saline) solution, individually identified, stored in liquid nitrogen and transported to the Animal Biotechnology Laboratory (LABTEC) at the Department of Animal Science, UFV.

Each conceptus was entirely and separately macerated in liquid nitrogen. Three embryos from 3 CON gilts and three embryos from 3 ARG gilts at 25 days of gestation and three fetuses from 3 CON gilts and three fetuses from 3 ARG gilts at 35 days of gestation were randomly chosen for RNA-seq analysis, totaling 36 samples. Afterwards, these samples were transported in liquid nitrogen to the Animal Genetics Laboratory at the Embrapa Swine and Poultry Nacional Research Center, Concordia, SC, Brazil, for further RNA extraction and library preparation.

2.2 RNA extraction and library preparation

Total RNA extraction of the 25 and 35-days-old conceptuses (100 mg) from CON and ARG females was performed with TRIzol® (1 mL) (Invitrogen, San Diego, USA), as described in Teixeira et al. (2019). Afterwards, RNA quantification was performed in QUBIT fluorimeter (Thermo Scientific, Waltham, USA). RNA integrity was determined in 1.0% agarose gel and in Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, USA), in which samples with RNA integrity number (RIN) score higher than eight were used for library preparation. Due to low RIN score (RIN < 8), one fetus sample was excluded from subsequent analysis. The 35 remaining conceptuses samples (nine CON and nine ARG at 25 days and nine CON and eight ARG at 35 days) had their total amount of extracted and purified were submitted to RNA-seq library preparation using the TruSeq Stranded mRNA

Library Prep Kit (Illumina, Inc., San Diego, USA), followed by purification of the poly-A tail using 2 µg of total RNA, according to the manufacturer´s recommendations.

2.3 Sequencing, quality control and assembly

The libraries were sequenced in Illumina HiSeq2500 (Illumina, Inc., San Diego, EUA), with a 2x100bp paired-end protocol, at the Functional Genomics Center, ESALQ, *Universidade de São Paulo*, Piracicaba, SP, Brazil. The sequence data have been submitted to the GenBank databases under accession number PRJNA576701 and the BAQCOM pipeline (https://github.com/hanielcedraz/BAQCOM) (BAQCOM: Bioinformatics Analysis for Quality Control and Mapping) was used to perform quality control (QC), mapping and reads counting. BAQCOM uses the softwares Trimmomatic (Bolger, Lohse, & Usadel, 2014) version 0.38, HISAT2 (Kim, Langmead, & Salzberg, 2015) version 0.11.2, and HTseq-count (Anders, Pyl, & Huber, 2015) version 0.11.2 to analyze RNA-seq data, as described in Teixeira et al. (2019). In these analyses, only reads with Phred quality ≥ 20 and length ≥ 70 pb were mapped against the pig reference genome (*Sus scrofa,* v. 11.1) and the reads counting for each feature was based on Ensembl annotation release 95 (https://www.ensembl.org). The identification of embryos (nine males and nine females) and fetuses' sex (10 males and seven females) in the RNA-seq analysis was performed according to Teixeira et al. (2019).

2.4 Differential expression analysis

The *limma* package (Ritchie et al., 2015) in R software (R Core Team, 2018) was used for data analysis and identification of differentially expressed (DE) genes. In these analyses, a mixed model with a random nested effect of diets (CON or ARG) within females was used, since conceptuses were considered false replicates of the females for each diet. The sex effect was considered as an adjustment factor in the statistical model.

Genes with positive and negative log fold-change (FC) values were considered, respectively, down and up-regulated in ARG conceptuses, according to the following statistical model:

$$Y_{ijk} = \mu + S_i + D_j + d\{D_{(j)}\}_k + \varepsilon_{ijk}$$

wherein: Y_{ijk} is the logarithm of the reads counts per million; μ is the trait general mean; S_i is the sex effect of i-th conceptus; D_j is the effect of the j-th diet (CON or ARG); $d\{D_{(j)}\}_k$ is the random effect of j-th diet nested within k-th female and ε_{ijk} is the random error.

Significance threshold for DE genes was set at adjusted *P-value*≤0.05 after multiple correction tests to reduce type I error, following the Benjamini and Hochberg method (Benjamini & Hochberg, 1995). The DE genes with unknown names in Ensembl database had their nucleotide sequences aligned against the NCBI database using the BLAST tool nucleotide and Blastx programs (Camacho et al., 2009).

2.5 Gene ontology and network analyses

The DE genes were used as input list and all the identified expressed genes were set as background list in DAVID tool 6.8 (Huang et al., 2007) to identify the

gene ontology (GO) terms. The significant GO terms (adjusted *P-value*≤0.05) in GOTERM_BP_ALL category were used as input on REVIGO (Reduce and Visualize Gene Ontology) web server (Supek, Bošnjak, Škunca, & Šmuc, 2011) to summarize and remove redundant GO terms. The output of REVIGO analysis was used as input on Cytoscape (Shannon et al., 2003), version 3.7.2, to perform the gene network analysis, with each gene associated with the correspondent pathway.

3. RESULTS

3.1 RNA-seq data

An average of 15.3 and 14.4 million reads/sample (2x100 bp) were sequenced for embryos (n=18) and fetuses (n=17), respectively. After the data QC, 13.6 and 12.9 million reads were kept for embryos and fetuses, respectively. More than 98% reads were mapped against the pig reference genome (*Sus scrofa*, v.11.1), in which 83.8% and 80.3% of the reads were counted within genes for embryos and fetuses samples, respectively.

3.2 Differentially expressed genes

A total of 14,442 and 14,701 transcripts were obtained in the transcriptome of embryos and fetuses, respectively.

Eight genes were DE between CON and ARG embryos groups: Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 1 (*ND1*), Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 2 *(ND2),* Mitochondrially Encoded Cytochrome B (*CYTB*), Mitochondrially Encoded Cytochrome C Oxidase II (*COX2*), Mitochondrially Encoded ATP Synthase

Membrane Subunit 8 (*ATP8*), ENSSSCG00000018063 (*rRNA-MT* type), *ENSSSCG00000035520* (novel gene) and Cytochrome P450 Family 1 Subfamily A Member 1 (*CYP1A1*). The transcripts *ND1*, *ND2, CYTB, COX2*, *ATP8* and *rRNA-MT type* were down-regulated in ARG embryos, while *ENSSSCG00000035520* and *CYP1A1* were up-regulated in this group compared to CON embryos (Table 1).

Table 1: Differentially expressed genes between control and arginine embryos from RNA-sequencing data.

Ensembl gene ID	Gene symbol	Description	logFC[1]	Adj. *P*-value[2]
ENSSSCG00000018065	*ND1*	Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 1	1.63	3.14E-09
ENSSSCG00000018069	*ND2*	Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 2	0.70	2.29E-02
ENSSSCG00000018094	*CYTB*	Mitochondrially Encoded Cytochrome B	0.51	9.69E-03
ENSSSCG00000018078	*COX2*	Mitochondrially Encoded Cytochrome C Oxidase II	0.75	1.65E-03
ENSSSCG00000018080	*ATP8*	ATP synthase F0 subunit 8	5.58	4.49E-10
ENSSSCG00000018063	*rRNA-MT type*	type mitochondrial ribosomal RNA	1.49	9.69E-03
ENSSSCG00000035520	*ENSSSCG00000035520*	novel gene	-8.77	1.04E-19
ENSSSCG00000001906	*CYP1A1*	Cytochrome P450 Family 1 Subfamily A Member 1	-1.75	4.95E-02

[1]log fold-change; [2]Adjusted *P*-value after multiple correction tests to reduce type I error; significance threshold: adjusted *P*-value≤0.05.

The DE *rRNA-MT type* gene and *ENSSSCG00000035520* uncharacterized protein had their nucleotide sequences aligned against the NCBI database using the BLAST tool nucleotide and BLASTX programs, respectively. The *rRNA-MT* type showed 100% alignment with the mitochondrial *16S rRNA Sus scrofa* gene (GenBank: AF034253.1), according to *Sus scrofa* mitochondrial genome (C. S. Lin et al., 1998) and the uncharacterized protein had 23.5% alignment with *Sus scrofa* NADH dehydrogenase subunit 1 (protein_ID: AGG55309.1). Due to low coverage (23.5%) with NADH dehydrogenase subunit 1 from Ensembl database, the *ENSSSCG00000035520* gene was not fully explored in the present study.

Regarding the fetuses, there were no DE genes on global transcriptome between CON and ARG diets.

3.3 Gene ontology and gene network

Using the DAVID tool, 48 GO terms were identified, of which 35 were significant associated with *ND1*, *ND2*, *CYTB*, *COX2*, *ATP8* and *CYP1A1* genes between CON and ARG embryos (Table 2). The *rRNA type,* similar to 16*S rRNA* gene, was not associated with GO terms, probably because functional annotation data for non-coding RNA are less pronounced than coding RNA (Conesa et al., 2016). The significant GO terms were grouped into 11 GO terms by the REVIGO web server (Table 3). The *ND1*, *ND2*, *CYTB*, *COX2*, and *ATP8* genes shared the biological processes (BP) of oxidation reduction, cellular respiration, and carbohydrate derivative, organophosphate, purine-containing compound, nucleoside triphosphate and glycosyl compound metabolic processes (Figure 1).

These genes also shared the BP of metabolism of organonitrogen compounds with the *CYP1A1* gene (Figure 1).

Table 2: Gene ontology from differentially expressed genes between control and arginine embryos, using DAVID tool

Term[1]	Pathways	Count[2]	Adj. *P*-value[3]
GO:0009141	nucleoside triphosphate metabolic process	4	1.11E-03
GO:0009123	nucleoside monophosphate metabolic process	4	1.19E-03
GO:0009161	ribonucleoside monophosphate metabolic process	4	1.26E-03
GO:0006091	generation of precursor metabolites and energy	4	1.30E-03
GO:0009167	purine ribonucleoside monophosphate metabolic process	4	1.35E-03
GO:0009126	purine nucleoside monophosphate metabolic process	4	1.35E-03
GO:0009119	ribonucleoside metabolic process	4	1.38E-03
GO:0042278	purine nucleoside metabolic process	4	1.38E-03
GO:0046128	purine ribonucleoside metabolic process	4	1.45E-03
GO:0009116	nucleoside metabolic process	4	1.51E-03
GO:0009144	purine nucleoside triphosphate metabolic process	4	1.58E-03
GO:1901657	glycosyl compound metabolic process	4	1.73E-03
GO:0009259	ribonucleotide metabolic process	4	1.80E-03
GO:0006163	purine nucleotide metabolic process	4	1.85E-03
GO:0019693	ribose phosphate metabolic process	4	1.85E-03
GO:0042773	ATP synthesis coupled electron transport	3	1.91E-03
GO:0009199	ribonucleoside triphosphate metabolic process	4	1.92E-03
GO:0009150	purine ribonucleotide metabolic process	4	1.93E-03
GO:0045333	cellular respiration	4	1.97E-03
GO:0006119	oxidative phosphorylation	3	2.01E-03
GO:0072521	purine-containing compound metabolic process	4	2.02E-03
GO:0022904	respiratory electron transport chain	3	2.31E-03
GO:0009205	purine ribonucleoside triphosphate metabolic process	4	2.43E-03
GO:0015980	energy derivation by oxidation of organic compounds	4	2.43E-03
GO:0022900	electron transport chain	3	2.59E-03
GO:0046034	ATP metabolic process	4	2.64E-03
GO:0009117	nucleotide metabolic process	4	2.78E-03
GO:0044710	single-organism metabolic process	6	2.81E-03
GO:0006753	nucleoside phosphate metabolic process	4	2.81E-03
GO:0055114	oxidation-reduction process	4	2.85E-03

GO:0055086	nucleobase-containing small molecule metabolic process	4	3.21E-03
GO:1901564	organonitrogen compound metabolic process	5	5.97E-03
GO:0019637	organophosphate metabolic process	4	8.90E-03
GO:1901135	carbohydrate derivative metabolic process	4	1.55E-02

[1]Gene Ontology (GO) Term_Biological Processes (BP); [2]Number of genes in GO term; [3]Adjusted *P*-value after multiple correction tests to reduce type I error; significance threshold: adjusted *P*-value≤0.05.

Table 3: Gene Ontology terms and pathways associated with differentially expressed genes between control and arginine embryos found using REVIGO tool

Term[1]	Pathways
GO:0009141	nucleoside triphosphate metabolic process
GO:1901135	carbohydrate derivative metabolic process
GO:0006091	generation of precursor metabolites and energy
GO:0044710	single-organism metabolic process
GO:1901564	organonitrogen compound metabolic process
GO:1901657	glycosyl compound metabolic process
GO:0045333	cellular respiration
GO:0072521	purine-containing compound metabolic process
GO:0055114	oxidation-reduction process
GO:0055086	nucleobase-containing small molecule metabolic process
GO:0019637	organophosphate metabolic process

[1]GO: Gene Ontology

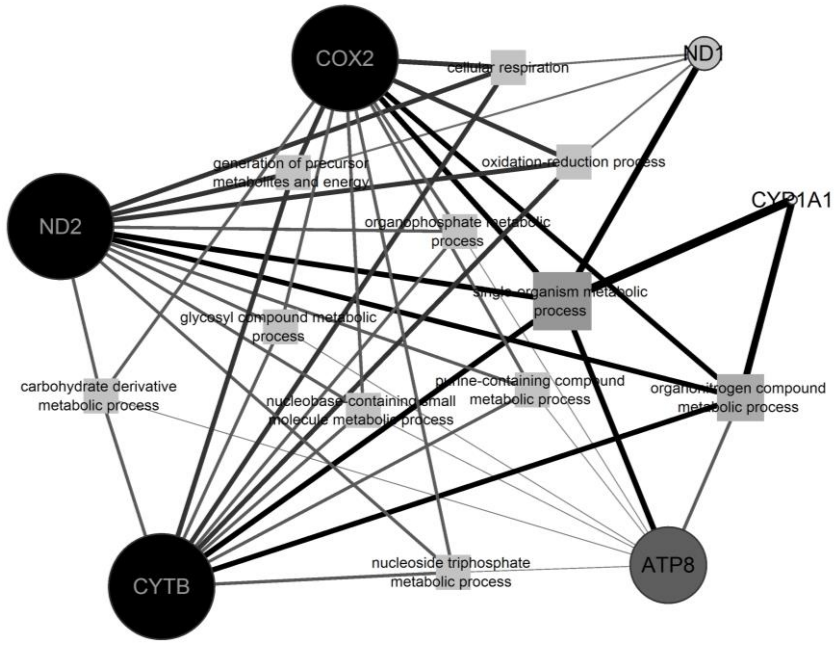


Figure 1: Gene network showing connections between differentially expressed genes (circles) and biological processes (squares). The nodes size and gray scale colors indicate the number of directed edges and the neighborhood connectivity, respectively (small size and light gray represent low values). The edge colors indicate the betweenness of the edges (light gray represent low values). ATP8, ATP synthase F0 subunit 8; COX2, mitochondrially encoded cytochrome c oxidase II; CYP1A1, cytochrome P450 family 1 subfamily A member 1; CYTB, mitochondrially encoded cytochrome b; ND1, mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 1; ND2, mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 2.

4. DISCUSSION

Arginine is a functional amino acid, serving as a substrate for the syntheses of proteins and bioactive molecules, such as NO and polyamines, which play a key role in the development of conceptuses and placental membranes (Wu et al., 2009, 2004). Due to its versatility, arginine endogenous synthesis is insufficient and this amino acid must be provided from diets during gestation. However, even knowing that maternal nutrition might affect fetal programming, studies adressing the effects of dietary L-arginine supplementation on the regulation of gene expression in the prenatal growth of conceptuses are still scarce (Wu et al., 2013; Zhang et al., 2019).

In the present study, from global RNA-seq of embryos, we have observed that 1.0% L-arginine supplementation in early gestation of gilts changed the expression of genes related with energy metabolism, especially mitochondrial genes involved in the oxidative phosphorylation system (OXPHOS). The biological processes highlighted by the gene network, mainly the oxidation reduction and cellular respiration, reinforced the association of the mitochondrial genes with OXPHOS. This system is the most well-known energy metabolic pathway in the mitochondria, which couples nutrient oxidation to ATP (adenosine triphosphate) (Dumollard, Carroll, Duchen, Campbell, & Swann, 2009). Upon implantation, the

mitochondria mass significantly increases in embryos (Lima, Burstaller, Sanchez-nieto, & Rodríguez, 2018; H. Lin et al., 2015), considering their role in energy generation and production of essential metabolites for cell fate during organogenesis (Harvey, 2019; Lima et al., 2018).

The mitochondria have their own genome and encode 13 essential OXPHOS subunits: seven subunits of Complex I (ND1-6 and ND4L), one subunit of Complex III (CYTB), three subunits of Complex IV (COX1-3) and two subunits of Complex V (ATP6 and 8) (Cruz-Reyes & Gray, 2018). The complexes I, III and IV are important to create a proton gradient across the mitochondrial inner membrane, which is necessary for ATP synthase catalysis (complex V). The mitochondrial genome also encodes two ribosomal RNAs (rRNAs; 12S rRNA and 16S rRNA) and 22 transfer RNAs (tRNAs) (Cruz-Reyes & Gray, 2018; Yang et al., 2014). The 1.0% L-arginine supplementation of gilts resulted in down-regulation of messenger RNA (mRNA) of subunits involved in the transfer of electrons by complexes I (ND1 and ND2), III (CYTB), IV (COX2) and V (ATP8) and a mitochondrial rRNA similar to 16S rRNA in embryos. This joint down-regulation was expected due to the nature of the polycistronic transcription in mammalian mitochondria genome (Scheibye-Alsing, Cirera, Gilchrist, Fredholm, & Gorodkin, 2007).

Mitochondria are important dynamic organelles for cell survival and their function is associated with the control of energetic cell metabolism and signal transduction (Guo et al., 2018) in order to promote cell homeostasis (Mark A. Johnson et al., 2014). During disturbances in energy metabolism, caused by the increase in energy demand and/or decrease in mitochondrial energy supply,

mitochondria proliferation occurs to restore the cellular energy status, which leads to greater mitochondrial genes transcription (Wiesner et al., 1999). Furthermore, disturbances in amino acid homeostasis might up-regulate *ND1*, *ND2*, *ND3*, *COX3* and *ATP6/8* mRNA in embryonic kidney cells, suggesting that mitochondrial function stimulation may offer a way to inhibit nutrient-demanding anabolism that drives cellular proliferation (Mark A. Johnson et al., 2014). On the other hand, the greater energy supply by higher glucose levels or reduced insulin levels in diabetic individuals have shown down-regulation of OXPHOS genes and proteins expression (Hwang et al., 2010; Liang et al., 2018). Similarly, in the present study, embryos from L-arginine supplemented gilts presented down-regulation of *ND1*, *ND2*, *CYTB*, *COX2* and *ATP8* OXPHOS genes.

During intra-uterine development, impairment of placental blood flow results in fetal under nutrition (Godfrey & Barker, 2001). Consequently, fetuses may reduce their metabolic dependence on glucose, increase oxidation of other energetic substrates, including amino acids, and reduce the concentration of anabolic hormones as insulin-like growth factor I (IGF-1) (Godfrey & Barker, 2001). In this context, our research group has observed increase of *IGF-1* mRNA expression in embryos from dietary L-arginine supplemented gilts (Costa et al., 2019). Arginine is a substrate for NO synthesis, the main endothelium-derived relaxing factor involved in the supply of respiratory substrates to mitochondria (Nisoli, Cozzi, & Carruba, 2008). In this way, Costa et al. (2019) also observed a trend toward a greater concentration of arginine in the blood plasma of ARG gilts at 25 days of gestation, which could result in greater NO production (Bird et al.,

2003). Considering that NO is involved in increasing the availability of nutrients and oxygen to conceptuses through control of blood flow (Bird et al., 2003; Wu et al., 2008), the down-regulation of OXPHOS genes in embryos from ARG gilts may suggest that these embryos have received greater supply of nutrients and energy than embryos from CON gilts, consequently reducing ATP production by OXPHOS. In addition, since alterations in nutrition may impact progeny metabolism (Wu et al., 2004) and considering that embryonic transcriptional program is strictly regulated (Campbell et al., 2005), the lower OXPHOS genes expression in embryos from ARG gilts may represent an adaptive metabolic response mediated by NO effects.

Throughout embryonic development, the ATP cellular requirement increases, resulting in greater mitochondrial activity (Balaban, 1990). Insufficient ATP production may lead to cell apoptosis, indicating reduction of cell viability (Brown, 2001; Dai et al., 2013). Although the lower expression of OXPHOS genes, especially the *ATP8*, may result in lower ATP synthesis, we did not identify changes in expression of apoptosis genes between CON and ARG embryos. Additionally, a previous study from our research group did not report differences between CON and ARG embryos for *BAX*, *BCL2* and *CASP3* apoptosis genes (Costa et al., 2019). Therefore, the reduction in mitochondria activity did not seem harmful for pig embryos development in the current condition.

Besides increasing ATP synthesis, greater mitochondrial activity may result in increased ROS (reactive oxygen species) production, enhancing lipid peroxidation, protein oxidation, nuclear DNA strand breaks and mitochondrial alteration (Guérin, El Mouatassim, & Ménézo, 2001; Ufer, Wang, Borchert, Heydeck, & Kuhn, 2010).

After yolk-sac regression, utero-placental and embryonic circulation systems develop, exposing embryos to greater oxygen concentrations, which makes the ROS production a more disturbing process (Ufer et al., 2010). In this way, according to Guérin et al. (2001), early embryos can change their own metabolism from more “oxidative” to more “anaerobic” and this may reduce mitochondria ROS production (De Boo et al., 2005; Dumollard et al., 2009; Houghton, 2003). Manser et al. (2004) showed that lower oxygen consumption is associated with greater developmental competence of blastocyst and it might be advantageous for porcine embryonic development *in vitro* (Macháty, Thompson, Abeydeera, Day, & Prather, 2001). Among OXPHOS complexes, complexes I and III are the most implicated in ROS generation (Hea et al., 2017). Then, since OXPHOS activity is closely related to the levels of gene expression encoding subunits of these respiratory complexes (Weller et al., 2013), a consequence of lower mRNA expression of complexes subunits I (*ND1* and *ND2*) and III (*CYTB*) in ARG embryos may be to avoid increased ROS production, which may ensure a satisfactory competence of embryos development rather than impairment on energy production.

Two rRNA are characterized and encoded by the mammalian mitochondrial genome: 12S rRNA and 16S rRNA (Yang et al., 2014). In the current study, we found a sequence similar to *Sus scrofa* mitochondrial rRNA 16 subunit (*mt-rRNA type*). The *16S rRNA* gene is a member of the 23S-like family of large subunit ribosomal RNAs (Burk, Douzery, & Springer, 2002) and is necessary for translation of mRNAs into mitochondrial proteins (Yang et al., 2014). Therefore, since mammalian mitochondria genome presents polycistronic nature and mitochondrial

ribosome mediates protein synthesis in the mitochondria, the down-regulation of *mt-rRNA type* is in agreement with the reduced mRNA levels of *ND1*, *ND2*, *CYTB*, *COX2* and *ATP8* OXPHOS genes in ARG embryos.

The placenta is an important organ during pregnancy, required for the transport of nutrients, oxygen and hormones from the mother to the fetuses and for the removal of waste products from the embryonic side (Hafez & E.S.E, 2000). Other components, such as xenobiotics, are also transferred across the placenta according to their physicochemical properties (Pacifici & Nottoli, 1995) and may act as chemical teratogens (Tayeboon et al., 2015). In this context, CYP1A1, encoded by *CYP1A1* gene, is known as a molecular marker that mediates the response of an organism to environmental challenges (Choudhary, Jansson, Schenkman, Sarfarazi, & Stoilov, 2003), since it is involved in detoxification of potentially toxic substances that might pass from maternal blood to the developing embryo and fetus (Campbell et al., 2005; Choudhary et al., 2003). In the present study, *CYP1A1* mRNA expression was greater in ARG embryos and this gene was associated with the biological process of nitrogenous compounds metabolism in the gene network. This result can be explained by increased supply of nitrogenous compounds or other organic metabolites from ARG mother to embryos mediated by arginine metabolites that increase blood flow, which may have led to the greater production of toxic or teratogen substances that might impair embryonic development (Choudhary et al., 2003). Consequently, embryos may activate mechanisms, such as increased production of CYP1A1 protein, to metabolize

those foreign compounds that can act as chemical teratogens (Tayeboon et al., 2015).

Although the results obtained in the present study suggest that the modulation of embryonic metabolism occurs in response to increased placental energetic supply and, consequently, increased availability of metabolites from mother to embryos, future studies should be performed to help clarifying the relationship between nutrient flow and metabolic activity in embryos, since several studies address the metabolism in the pre-implantation stage (Dumollard et al., 2009; Guérin et al., 2001; Leese, 2002; Liang et al., 2018; Macháty, Day, & Prather, 2005; Macháty et al., 2001; Manser et al., 2004).

Regarding the pig fetuses' development, the amino acid accretion, which represents minimal requirements of amino acids by the fetus, quickly increases with gestation and following fetal growth (Wu et al., 1998). It is mainly related to the quick skeletal muscle development, since the primary myogenic wave begins between 25-50 days of gestation and the secondary myogenic wave occurs between 50 and 90 days of gestation (Ji et al., 2017). In this context, Bérard & Bee (2010) reported that L-arginine supplementation of gilts between 14 and 28 days of gestation positively affected the primary phase of myofiber formation, possibly due to the improvement of the uterine environment, such as increased placental angiogenesis, resulting in enhanced supply of nutrients to developing fetuses. Therefore, a greater nutrient supply to fetuses around 30 days-old mediated by the effects of gilts L-arginine supplementation would be important to program the fetal

development, since the requirements for muscle development increase in this period.

Nevertheless, L-arginine supplementation of pregnant gilts during 33 days of gestation did not affect the global fetuses' transcriptome at 35 days-old in the current study. The lack of differences between global transcriptome of CON and ARG fetuses may be due to some main factors. Firstly, L-arginine supplementation during 33 days may have led to its greater metabolization by arginase activity, reducing its availability for NO production by endothelial NO synthases (Dioguardi, 2011). In this context, data from a previous study of our research group showed that at 35 days of gestation, the arginine concentration in the blood plasma of L-arginine supplemented gilts tended to be reduced compared with non-supplemented females (Costa, et al., 2019). Therefore, although arginine supplementation in early gestation may be beneficial (Bérard & Bee, 2010; Bass, Bradley, Johnson, Boyd, et al., 2017; Palencia et al., 2017), the optimum period and duration of supplementation still need to be established (Costa et al., 2019; Palencia et al., 2017; Zhang et al., 2019), which justifies different research paths to determine L-arginine effects on pigs prenatal development.

Secondly, L-arginine metabolites can also be recruited by other metabolic pathways rather than developing fetuses, since there are several other metabolic pathways in which this amino acid is involved, e.g. mammary gland growth. In this context, Pau & Milner, (1982) reported that deficiencies in dietary arginine during gestation in rats reduced the number of mammary cells and the gland functional

activity, suggesting that arginine is an essential amino acid for optimal mammary growth during pregnancy.

5. CONCLUSION

1.0% L-arginine supplementation of gilts until 25 days of gestation changed global embryos' transcriptome through down-regulation of OXPHOS and detoxification genes, suggesting a regulation of energy metabolism in response to greater supply of nutrients in embryos and reinforcing an adjustment of metabolism in order to ensure homeostasis during embryos' development. On the other hand, the 35 days old fetuses' transcriptome was not affected by L-arginine, providing support for further discussions regarding the optimal period of supplementation in order to understand its impacts on swine prenatal development.

ACKNOWLEDGMENTS

This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Grant # 423558/2018-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Grant # PROEX 0577/2018 and 320.020.170.11pq), Instituto Nacional de Ciência e Tecnologia (INCT-CA, Grant # 465377/2014-9) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Grant # CVZ APQ 03434/16). MCL, LLC and SEFG are CNPq fellows. The authors also thank Professor Gordon K. Smyth, from Walter and Eliza Hall Institute of Medical Research, for his valuable support with the analyses.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ORCID

Susana A. Teixeira https://orcid.org/0000-0002-2232-1417
Adriana M. G. Ibelli https://orcid.org/0000-0003-0626-8879
Maurício E. Cantão https://orcid.org/0000-0002-5988-6542
Haniel C. de Oliveira https://orcid.org/0000-0001-7317-2004
Daniele B. D. Marques https://orcid.org/0000-0001-7080-7465
Mônica C. Ledur https://orcid.org/0000-0002-2645-4735
Jane de O. Peixoto https://orcid.org/0000-0002-0644-3684
Karine A. Costa https://orcid.org/0000-0001-6227-4670
Luiz L. Coutinho https://orcid.org/0000-0002-7266-8881
Alysson Saraiva https://orcid.org/0000-0002-1603-0955
Simone E. F. Guimarães https://orcid.org/0000-0003-3704-8131

REFERENCES

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. doi:10.1093/bioinformatics/btu638

Balaban, R. S. (1990). Regulation of oxidative phosphorylation in the mammalian cell. *American Physiology Society*, 258(3), C377–C389. doi:10.1152/ajpcell.1990.258.3.c377

BAQCOM: Bioinformatics Analysis for Quality Control and Mapping. (n.d.). Retrieved 15 February 2019, from https://github.com/hanielcedraz/BAQCOM

Bass, B. E., Bradley, C. L., Johnson, Z. B., Boyd, R. D., Usry, J. L., Maxwell, C. V, & Frank, J. W. (2017). Influence of dietary L-arginine supplementation of sows during late pregnancy on piglet birth weight and sow and litter performance during lactation. *American Society of Animal Science*, 95(5), 248–256. doi:10.2527/jas2016.0986

Bazer, F. W., Spencer, T. E., Johnson, G. A., Burghardt, R. C., & Wu, G. (2009). Comparative aspects of implantation. *Reproduction*, 138(2), 195–209. doi:10.1530/rep-09-0158

Benjamini, Y., & Hochberg, Y. (1995). Controllin the false discovery rate: a practical and powerful approch to multiple testing. *Journal of the Royal Statistic Society*, 57, 289–300.

Bérard, J., & Bee, G. (2010). Effects of dietary L-arginine supplementation to gilts during early gestation on foetal survival, growth and myofiber formation. *Animal*, 4(10), 1680–1687. doi:10.1017/S1751731110000881

Bird, I. M., Zhang, L., & Magness, R. R. (2003). Possible mechanisms underlying pregnancy-induced changes in uterine artery endothelial function. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 13–28. doi: 10.1152/ajpregu.00108.2002

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

doi:10.1093/bioinformatics/btu170
Brown, G. C. (2001). Regulation of mitochondrial respiration by nitric oxide inhibition of cytochrome c oxidase. *Biochimica et Biophysica Acta - Bioenergetics*, 1504(1), 46–57. doi:10.1016/S0005-2728(00)00238-3
Burk, A., Douzery, E. J. P., & Springer, M. S. (2002). The secondary structure of mammalian mitochondrial 16S rRNA molecules: Refinements based on a comparative phylogenetic approach. *Journal of Mammalian Evolution*, 9(3), 225–252. doi:10.1023/A:1022649516930
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST +: architecture and applications. *BMC Bioinformatics*, 9, 1–9. doi:10.1186/1471-2105-10-421
Campbell, S. J., Henderson, C. J., Anthony, D. C., Davidson, D., Clark, A. J., & Wolf, C. R. (2005). The Murine Cyp1a1 Gene Is Expressed in a Restricted Spatial and Temporal Pattern during Embryonic Development. *The American Society for Biochemistry and Molecular Biology*, 280(7), 5828–5835. doi:10.1074/jbc.M412899200
Choudhary, D., Jansson, I., Schenkman, J. B., Sarfarazi, M., & Stoilov, I. (2003). Comparative expression profiling of 40 mouse cytochrome P450 genes in embryonic and adult tissues. *Archives of Biochemistry and Biophysics*, 414, 91–100. doi:10.1016/S0003-9861(03)00174-7
Conesa, A., Madrigal, P., Tarazona, S., Gomez-cabrero, D., Cervera, A., Mcpherson, A., … Zhang, X. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17, 1–19. doi:10.1186/s13059-016-0881-8
Costa, K. A., Saraiva, A., Guimaraes, J. D., Marques, D. B. D., Machado-neves, M., Reis, L. M. B., … Guimarães, S. E. F. (2019). Dietary L-arginine supplementation during early gestation of gilts affects conceptuses development. *Theriogenology*, 140, 62–71. doi:10.1016/j.theriogenology.2019.08.018
Cruz-Reyes, J., & Gray, M. W. (2018). *RNA Metabolism in Mitochondria*. (J. Cruz-Reyes & M.W. Gray,Eds.). Philadelphia: Springer International Publishing AG, part of Springer Nature 2018. doi:https://doi:10.1007/978-3-319-78190-7
Dai, Z., Wu, Z., Yang, Y., Wang, J., Satterfield, M. C., Meininger, C. J., … Wu, G. (2013). Nitric oxide and energy metabolism in mammals. *BioFactors*, 39(4), 383–391. doi:10.1002/biof.1099
De Boo, H. A., Van Zijl, P. L., Smith, D. E. C., Kulik, W., Lafeber, H. N., & Harding, J. E. (2005). Arginine and mixed amino acids increase protein accretion in the growth-restricted and normal ovine fetus by different mechanisms. *Pediatric Research*, 58(2), 270–277. doi:10.1203/01.PDR.0000169977.48609.55
Dioguardi, F. S. (2011). To give or not to give? Lessons from the arginine paradox. *Journal of Nutrigenetics and Nutrigenomics*, 4(2), 90–98. doi:10.1159/000327777
Dumollard, R., Carroll, J., Duchen, M. R., Campbell, K., & Swann, K. (2009). Mitochondrial function and redox state in mammalian embryos. *Seminars in Cell and Developmental Biology*, 20(3), 346–353. doi:10.1016/j.semcdb.2008.12.013
Gao, K., Jiang, Z., Lin, Y., Zheng, C., Zhou, G., Chen, F., … Wu, G. (2012). Dietary l-arginine supplementation enhances placental growth and reproductive performance in sows. *Amino Acids*, 42(6), 2207–2214. doi:10.1007/s00726-011-0960-9

Godfrey, K. M., & Barker, D. J. P. (2001). Fetal programming and adult health. *Public Health Nutrition*, 4, 611–624. doi:10.1079/PHN2001145

Guérin, P., El Mouatassim, S., & Ménézo, Y. (2001). Oxidative stress and protection against reactive oxygen species in the pre-implantation embryo and its surroundings. *Human Reproduction Update*, 7(2), 175–189. doi:10.1093/humupd/7.2.175

Guo, Y. X., Nie, H. T., Xu, C. J., Zhang, G. M., Sun, L. W., Zhang, T. T., … Wang, F. (2018). Effects of nutrient restriction and arginine treatment on oxidative stress in the ovarian tissue of ewes during the luteal phase. *Theriogenology*, 113, 127–136. doi:10.1016/j.theriogenology.2018.02.016

Hafez, B., & E.S.E, H. (2000). *Reproduction in farm animals* (7nd ed), South Carolina, USA.

Harvey, A. J. (2019). Mitochondria in early development: linking the microenvironment, metabolism and the epigenome. *Reproduction*, 159–179. doi:10.1530/REP -18-0431

Hea, L., Hea, T., Farrarb, S., Jia, L., Liua, T., & Maa, X. (2017). Antioxidants Maintain Cellular Redox Homeostasis by Elimination of Reactive Oxygen Species. *Cellular Physiology and Biochemistry*, 532–553. doi:10.1159/000485089

Houghton, F. D. (2003). Non-invasive amino acid turnover predicts human embryo developmental capacity. *Human Reproduction*, 18(8), 1756–1757. doi:10.1093/humrep/deg389

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., … Lempicki, R. A. (2007). DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research*, 35(SUPPL.2), 169–175. doi:10.1093/nar/gkm415

Hwang, H., Bowen, B. P., Lefort, N., Flynn, C. R., De Filippis, E. A., Roberts, C., … Mandarino, L. J. (2010). Proteomics analysis of human skeletal muscle reveals novel abnormalities in obesity and type 2 diabetes. *Diabetes*, 59(1), 33–42. doi:10.2337/db09-0214

Ji, Y., Wu, Z., Dai, Z., Wang, X., Li, J., Wang, B., & Wu, G. (2017). Fetal and neonatal programming of postnatal growth and feed efficiency in swine. *Journal of Animal Science and Biotechnology*, 8(1), 1–15. doi:10.1186/s40104-017-0173-5

Keshet, R., & Erez, A. (2018). Arginine and the metabolic regulation of nitric oxide synthesis in cancer. *Disease Models & Mechanisms*, 11(8). doi:10.1242/dmm.033332

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*, 12(4), 357–360. doi:10.1038/nmeth.3317

Kwon, E. J., & Kim, Y. J. (2017). What is fetal programming?: A lifetime health is under the control of in utero health. *Obstetrics and Gynecology Science*, 60(6), 506–519. doi: 10.5468/ogs.2017.60.6.506

Leese, H. J. (2002). Quiet please, do not disturb: A hypothesis of embryo metabolism and viability. *BioEssays*, 24(9), 845–849. doi:10.1002/bies.10137

Li, X., Bazer, F. W., Johnson, G. A., Burghardt, R. C., Erikson, D. W., Frank, J. W., … Wu, G. (2010). Dietary Supplementation with 0.8% L-Arginine between Days 0 and 25 of Gestation Reduces Litter Size in Gilts. *The Journal of Nutrition*, 140(6), 1111–1116. doi:10.3945/jn.110.121350

Li, X., Bazer, F. W., Johnson, G. A., Burghardt, R. C., Frank, J. W., Dai, Z., … Wu, G. (2014). Dietary supplementation with l-arginine between days 14 and 25 of gestation enhances embryonic development and survival in gilts. *Amino Acids*, 46(2), 375–384. doi:10.1007/s00726-013-1626-6

Liang, Q., Qu, Z., Liang, Y., Feng, Q., Niu, X., Bai, T., … Adelson, D. L. (2018). Zuo Gui Wan Alters Expression of Energy Metabolism Genes and Prevents Cell Death in High-Glucose Loaded Mouse Embryos. *Hindawi*. doi: 10.1155/2018/2409471

Lima, A., Burstaller, J., Sanchez-nieto, J. M., & Rodríguez, T. A. (2018). The Mitochondria and the Regulation of Cell Fitness During Early Mammalian Development. *Current Topics in Developmental Biology*, 128, 339–363. doi:10.1016/bs.ctdb.2017.10.012

Lin, C. S., Liu, C. Y., Wu, H. T., Sun, Y. L., Chang, L. C., Yen, N. T., … Mao, S. J. T. (1998). SSCP analysis in the D-loop region of porcine mitochondrial DNA as confirmed by sequence diversity. *Journal of Animal Breeding and Genetics*, 115(1), 73–78. doi:10.1111/j.1439-0388.1998.tb00329.x

Lin, H., Wang, H., Wang, Y., Liu, C., Wang, C., & Guo, J. (2015). Transcriptomic Analysis of the Porcine Endometrium during Embryo Implantation. *Genes*, 1330–1346. doi:10.3390/genes6041330

Macháty, Z., Day, B. N., & Prather, R. S. (2005). Development of Early Porcine Embryos In Vitro and In Vivo1. *Biology of Reproduction*, 59(2), 451–455. doi:10.1095/biolreprod59.2.451

Macháty, Z., Thompson, J. G., Abeydeera, L. R., Day, B. N., & Prather, R. S. (2001). Inhibitors of mitochondrial ATP production at the time of compaction improve development of in vitro produced porcine embryos. *Molecular Reproduction and Development*, 58(1), 39–44. doi: 10.1002/1098-2795(200101)58:1<39::AID-MRD6>3.0.CO;2-B

Manser, R. C., Leese, H. J., & Houghton, F. D. (2004). Effect of Inhibiting Nitric Oxide Production on Mouse Preimplantation Embryo Development and Metabolism1. *Biology of Reproduction*, 71(2), 528–533. doi:10.1095/biolreprod.103.025742

Mark A. Johnson, Vidoni, S., Durigon, R., Pearce, S. F., Rorbach, J., He, J., … Spinazzola, A. (2014). Amino acid starvation has opposite effects on mitochondrial and cytosolic protein synthesis. *PLoS ONE*, 9(4). doi:10.1371/journal.pone.0093597

Mateo, R. D., Wu, G., Bazer, F. W., Park, J. C., Shinzato, I., & Kim, S. W. (2007). Dietary L-arginine supplementation enhances the reproductive performance of gilts. *The Journal of Nutrition*, 137(3), 652–656. doi:137/3/652

Nisoli, E., Cozzi, V., & Carruba, M. O. (2008). Amino Acids and Mitochondrial Biogenesis. *American Journal of Cardiology*, 101, 22–25. doi:10.1016/j.amjcard.2008.02.077

Pacifici, G. M., & Nottoli, R. (1995). Placental Transfer of Drugs Administered to the Mother. *Cilinical Pharmacokinetic Concepts*, 28(3), 235–269. doi: 10.2165/00003088-199528030-00005

Palencia, J. Y. P., Lemes, M. A. G., Garbossa, C. A. P., Abreu, M. L. T., Pereira, L. J., & Zangeronimo, M. G. (2017). Arginine for gestating sows and foetal

development : A systematic review. *J Anim Physiol Anim Nutr*, 1–10. doi:10.1111/jpn.12679
Pau, M. Y., & Milner, J. A. (1982). Effect of arginine deficiency on mammary gland development in the rat. *Journal of Nutrition*, 112(10), 1827–1833. doi:10.1093/jn/112.10.1827
R Core Team. (2018). A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna.
Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), 1–13. doi:10.1093/nar/gkv007
Scheibye-Alsing, K., Cirera, S., Gilchrist, M. J., Fredholm, M., & Gorodkin, J. (2007). EST analysis on pig mitochondria reveal novel expression differences between developmental and adult tissues. *BMC Genomics*, 8, 1–13. doi:10.1186/1471-2164-8-367
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models. *Genome Research*, 13(22), 2498–2504. doi:10.1101/gr.1239303
Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, 6(7). doi:10.1371/journal.pone.0021800
Tayeboon, G. S., Osstad, S. N., Nasri, S., Nili-Ahmadabadi, A., Tavakoli, F., & Sabzevari, O. (2015). Effects of Teratogenic Drugs on CYP1A1 Activity in Differentiating Rat Embryo Cells. *Drug Res*, 65, 238–243. doi: 10.1055/s-0034-1370966
Teixeira, S. A., Ibelli, A. M. G., Cantão, M. E., Oliveira, H. C. De, Ledur, M. C., Peixoto, J. de O., … Guimarães, S. E. F. (2019). Sex Determination Using RNA-Sequencing Analyses in Early Prenatal Pig Development. *Genes*, 11. doi: 10.3390/genes10121010
Ufer, C., Wang, C. C., Borchert, A., Heydeck, D., & Kuhn, H. (2010). Redox Control in Mammalian Embryo Development. *Antioxidants & Redox Signaling*, 13(6), 833–875. doi:10.1089/ars.2009.3044
Wang, J., Feng, C., Liu, T., Shi, M., Wu, G., & Bazer, F. W. (2017). Physiological alterations associated with intrauterine growth restriction in fetal pigs: Causes and insights for nutritional optimization. *Molecular Reproduction and Development*, 84(9), 897–904. doi:10.1002/mrd.22842
Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, 57–63. doi: 10.1038/nrg2484
Weller, M. M. D. C. A., Alebrante, L., Campos, P. H. R. F., Saraiva, A., Silva, B. A. N., Donzele, J. L., … Guimarães, S. E. F. (2013). Effect of heat stress and feeding phosphorus levels on pig electron transport chain gene expression. *Animal*, 7(12), 1985–1993. doi:10.1017/s1751731113001535
Wiesner, R. J., Hornung, T. V, Garman, J. D., Clayton, D. A., Gorman, E. O., & Wallimann, T. (1999). Stimulation of Mitochondrial Gene Expression and Proliferation of Mitochondria Following Impairment of Cellular Energy Transfer by Inhibition of the Phosphocreatine Circuit in Rat Hearts. *Journal of Bioenergetics and Biomembranes,* 31(6), 559–567. doi: 10.1023/a:1005417011436

Wu, G. (2014). Dietary requirements of synthesizable amino acids by animals : a paradigm shift in protein nutrition. *Journal of Animal Science and Biotechnology*, 1–12. doi: 10.1186/2049-1891-5-34

Wu, G., Bazer, F. W., Cudd, T. a, Meininger, C. J., & Spencer, T. E. (2004). Recent Advances in Nutritional Sciences Maternal Nutrition and Fetal. *The Journal of Nutririon*, (13), 2169–2172.

Wu, G., Bazer, F. W., Datta, S., Gao, H., Johnson, G. A., Lassala, A., … Spencer, T. E. (2008). Intrauterine growth retardation in livestock : Implications , mechanisms and solutions. *Archiv Tierzucht, Dummerstrof*, 51, 4–10.

Wu, G., Bazer, F. W., Davis, T. A., Kim, S. W., Li, P., Marc Rhoads, J., … Yin, Y. (2009). Arginine metabolism and nutrition in growth, health and disease. *Amino Acids*, 37(1), 153–168. doi:10.1007/s00726-008-0210-y

Wu, G., Bazer, F. W., Johnson, G. A., & Hou, Y. (2018). Board-invited review: Arginine nutrition and metabolism in growing, gestating, and lactating swine. *Journal of Animal Science*, 96(12), 5035–5051. doi: 10.1093/jas/sky377

Wu, G., Bazer, F. W., Satterfield, M. C., Li, X., Wang, X., Johnson, G. A., … Wu, Z. (2013). Impacts of arginine nutrition on embryonic and fetal development in mammals. *Amino Acids*, 45(2), 241–256. doi:10.1007/s00726-013-1515-z

Wu, G., & Morris, S. M. (1998). Arginine metabolism: nitric oxide and beyond. *The Biochemical Journal*, 336, 1–17. doi: 10.1042/bj3360001.

Wu, G., Ott, T. L., Knabe, D. A., & Bazer, F. W. (1998). Amino Acid Composition of the Fetal Pig. *Nutrient Metabolism*, 1031–1038. doi: 10.1093/jn/129.5.1031

Yang, L., Tan, Z., Wang, D., Xue, L., Guan, M. X., Huang, T., & Li, R. (2014). Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports*, 4, 1–11. doi:10.1038/srep04089

Zhang, S., Heng, J., Song, H., Zhang, Y., Lin, X., Tian, M., … Guan, W. (2019). Role of Maternal Dietary Protein and Amino Acids on Fetal Programming, Early Neonatal Development, and Lactation in Swine. *Animals*, 9(1), 19. doi: 10.3390/ani9010019

CHAPTER 4

Transcription landscape of the developmental biology in pigs

ABSTRACT: During organogenesis, a tight transcriptional program is associated with the normal development of organs, tissues, and body structures. Since the trajectory of prenatal and postnatal developmental is usually programmed during prenatal life, studies that address the complete transcriptional landscape related to organogenesis are needed. Therefore, we aimed to evaluate the differentially expressed (DE) genes between embryos and fetuses from RNA-seq analysis using pig as a model organism. A total of 1,705 genes were DE between embryos at 25 days-old and fetuses at 35 days-old. Several of these genes enriched crucial GO terms related to developmental biology, including anatomical structure development, regulation of cell growth, negative regulation of development process and regulation of biologic quality. Among the top DE genes are *IBSP*, *COL6A6*, *HBE*, *HBZ*, *HBB*, *NEUROD6* genes, which were associated to events related to the physiology of embryonic to fetal developmental transition, such as ossification, cardiovascular system, skeletal muscle development, extracellular matrix organization, erythrocyte differentiation, and neuronal system. In addition, pathways related to myogenesis, neuronal development, cardiac and striated muscle contraction, were enriched for fetal development, reflecting the greater complexity of organs and body structures in this stage of development. On the other hand, the embryonic development was enriched by pathways mainly related to morphogenic signaling and cell interactions, which are crucial to the

transcriptional control at early prenatal development. Therefore, these findings unraveled several transcriptional changes associated with pig developmental biology, which can contribute to further prenatal issues, including the fetal program development as well as developmental anomalies.

Keywords: organogenesis, prenatal development, RNA-seq

_____

Article to be submitted

1. Introduction

The prenatal development in mammals is characterized by three main stages: the pre-implantation period, which comprise the fertilization to attachment of embryo into the uterine wall; the organogenesis, in which the main organs, body structures and tissues are formed; followed by its refinement and growth, characterizing the fetal period [1–3]. During embryogenesis, the development of complex and multicellular tissues and organs requires a well-orchestrated set of cellular events, involving cell-cell signaling and an intrinsic transcriptional program [4]. For this reason, the early prenatal development is highly susceptibility to developmental abnormalities, which can lead to impaired survival of conceptuses [3].

In pigs, a substantial increase of prenatal losses has been reported as a consequence of the uterine crowding, mainly in high prolific breeds [5,6]. On the other hand, the surviving conceptuses under uterine crowding conditions may present the intrauterine growth restriction (IUGR) [6], resulting in developmental biology impairments, such as reduced number of pancreatic islets, kidney glomeruli and muscle fibers, and ultimately, compromising the productivity and functionality of the affected progeny in postnatal life [7]. As many of these adverse effects on postnatal life are usually programmed in the prenatal development [6], the knowledge of the key genes that regulate biological events during prenatal development can provide a valuable support for studies addressing the causes and consequences of the developmental abnormalities, from use of marker genes and

their signaling pathways to investigate prenatal programming in the molecular context.

Several studies have reported the transcriptional changes during prenatal development using pig as model; however, most of them has been conducted in embryos or endometrium in the pre-implantation period [8–10], whereas others address the prenatal events mainly related to myogenesis from a limited set of genes obtained from microarray experiments [11,12]. Therefore, the complete background regarding the many events related to transcriptional program during the organogenesis of pig conceptuses remains poorly understood.

Based on this, in the present study, a high-throughput transcriptome analysis was used to identify the differentially expressed genes between embryos (at 25 days-old) and early fetuses (at 35 days-old) from which the main transcriptional changes were revealed, as well as the biological processes and pathways associated with developmental biology. The biological events described in this study provide valuable information to improve swine production and support studies on human prenatal development, since pig is an useful animal model for various studies addressing human genetic diseases and developmental biology, given the genetic and physiological similarity among them [13].

Material and Methods

Animal experiment

The experimental protocol has followed ethical principles in animal research (CONCEA, 2016) and was approved by the Ethical Committee on Animal Use of

Universidade Federal de Viçosa (UFV), Minas Gerais, Brazil [protocol no. 06/2017].

The present study is part of a whole experiment that was initially designed to understand the effects of maternal dietary L-arginine supplementation at 25 and 35 days of gestation in conceptuses development from transcriptome analyses [14,15]. Since the current study aimed to investigate the effects of development stages, instead of the effects of L-arginine supplementation, only data from embryos and fetuses from non-supplemented gilts (n=3 embryos from each three non-supplemented gilts at 25 days of gestation and n=3 fetuses from each three non-supplemented gilts at 35 days of gestation) were analyzed. Therefore, the experimental design is only briefly commented and additional details are available in *Costa et al.,* [15].

Briefly, 23 commercial gilts were inseminated 12 and 24 hours after the beginning of fourth estrus and the first insemination day was considered day zero of gestation. A total of 11 gilts were fed a control diet (CON) and 12 gilts were fed the CON diet supplemented with 1.0% L-arginine (ARG), using completely randomized experimental design with two diets (CON and ARG), and two gestational ages were considered (25 and 35 days of gestation). From 23 inseminated gilts, 20 became pregnant. A total of five females of CON and five females of ARG (at 25 days of gestation) and four CON gilts and six ARG gilts (at 35 days of gestation) were rendered unconscious using head-only electrical stunning (240V, 1.3A) and immediately exsanguinated. After slaughter, four conceptuses were collected per female at each gestational age. The conceptuses

were quickly washed with PBS (Phosphate Buffered Saline) solution, individually identified, stored in liquid nitrogen, and transported to the Animal Biotechnology Laboratory (LABTEC) at the Department of Animal Science, UFV. At the LABTEC, each conceptus was entirely and separately macerated in liquid nitrogen.

From these, three embryos from 3 CON gilts and three embryos from 3 ARG gilts at 25 days of gestation and three fetuses from 3 CON gilts and three fetuses from 3 ARG gilts at 35 days of gestation were randomly chosen for RNA-seq analysis, totalizing 36 samples, and transported in liquid nitrogen to the Animal Genetics Laboratory at the Embrapa Swine and Poultry Nacional Research Center, Concordia, SC, Brazil, for further RNA extraction and library preparation.

RNA extraction, library preparation and sequencing

These analyses were performed as previously described in *Teixeira et al.,*[16]. Briefly, 100 mg of whole macerated conceptuses were mixed to 1 mL Trizol® reagent (Invitrogen, San Diego, CA, USA) and the Qiagen RNeasy® Mini kit (Qiagen, Hilden, Germany) was used to total RNA isolation, following the manufacturer's instructions. After, the QUBIT fluorimeter (Thermo Scientific,Waltham, MA, USA) was used to quantify the extracted RNA and its integrity was determined in 1.0% agarose gel. In addition, the RNA integrity number (RIN) > 8 was evaluated using an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA). A total of 35 conceptuses samples were submitted to truSeq Stranded mRNA Library Prep Kit (Illumina, Inc., San Diego, CA, USA), followed by purification of the poly-A tail using 2 µg of total RNA starting material, according to the manufacturer´s recommendations. Sequencing was

performed in Illumina HiSeq2500 (Illumina, Inc.; San Diego CA, EUA), following the 2x100bp paired-end protocol, at the Functional Genomics Center, ESALQ, University of São Paulo, Piracicaba, SP, Brazil.

From the 35 sequenced samples, 18 conceptuses samples from the CON gilts (n=9 embryos at 25 days-old and n=9 fetuses at 35 days-old from CON gilts) were used in the present study to obtain an overview of transcriptional changes between the development of pig embryos and fetuses.

RNA-Sequencing data analysis

The sequence data were deposited in the Sequence Read Archive of the NCBI databases under Bioproject number PRJNA576701 and the 18 conceptuses samples used in the present study are under Biosample numbers SAMN13003040, SAMN13003041, SAMN13003042, SAMN13003043, SAMN13003044, SAMN13003045, SAMN13003046, SAMN13003047, SAMN13003048, SAMN13003049, SAMN13003050, SAMN13003051, SAMN13003052, SAMN13003053, SAMN13003054, SAMN13003055, SAMN13003056 and SAMN13003057.

Raw reads were quality trimmed with Trimmomatic tool [17], version 0.38, using the Phred quality ≥ 20 and length ≥ 70 pb as parameters. Then, the remaining sequence reads were aligned against the pig reference genome (*Sus scrofa,* v. 11.1) using Hisat software [18], version 0.11.2, and the read counting was performed with Htseq count software [19], version 0.11.2, using Ensembl annotation (release 95). All these analyses were performed following the BAQCOM

pipeline [20]. The conceptuses' sex was determined according to *Teixeira et al.,*[16].

The differentially expressed (DE) genes between fetuses and embryos were accessed using *limma* [21] package in R environment [R Core Team] and the sex effect was used as an adjustment factor in this analysis. Positive and negative log2 fold change (logFC) indicates genes up-regulated in fetuses and embryos, respectively. Genes were considered DE using a set of significance threshold at adjusted *p-value* for the false discovery rate (FDR) < 0.05, according to Benjamini-Hochberg method [23] and $|\log_2$ fold change (logFC)| > 0.5.

Functional Analyses

The functional annotation of DE genes was performed in the Database for Annotation, Visualization, and Integrated Discovery (DAVID) tool 6.8 (https://david.ncifcrf.gov/tools.jsp) [24], using all DE genes as input list and the *Sus scrofa* genome as the background list. The biological process (BP) category was considered as functionally enriched at FDR < 0.05 after Benjamini-Hochberg procedure [23] and submitted to ReviGO [25] to summarize and visualize the gene ontology (GO) terms. The Toppcluster [26] online tool (http://toppcluster.cchmc.org) was used to visualize enriched pathways (FDR < 0.05) among up-regulated genes in embryos (cluster 1) and up-regulated genes in fetuses (cluster 2). Finally, the visualization of functional enrichment from GO terms and network pathways were improved using Cytoscape [27], version 3.8.0.

Results

RNA-Seq data

From the whole transcriptome of conceptuses samples (n=9 embryos and n=9 fetuses), an average of 29.4 million reads/samples was generated. After the data quality control, an average of 26.1 million reads/samples was kept. Moreover, an average of 98.7% of the reads was mapped against the pig reference genome (*Sus scrofa*, v.11.1, Ensembl 95).

Differentially expressed genes and functional analyses

A total of 14,587 genes were expressed in conceptuses samples. A multiple dimensional scaling (MDS) analysis was performed from all these detected genes to evaluate the global transcriptome of the samples, which revealed an evident separation of the embryos and fetuses samples (Figure 1).

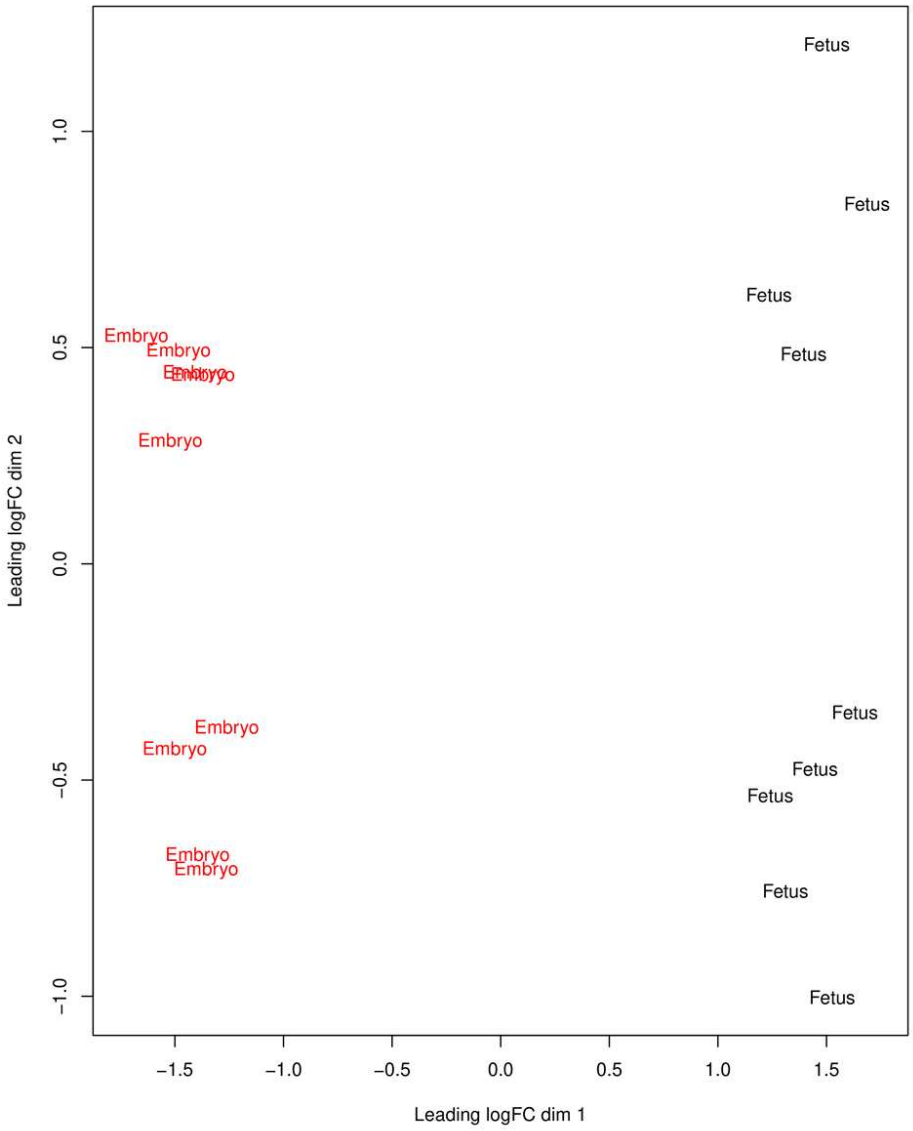


Figure 1: Multiple Dimensional Scale (MDS) plot of embryos samples at 25 days-old in red and fetuses samples at 35 days-old in black, based on all genes detected in pig prenatal transcriptome.

In DE analyses, a total of 1,705 genes were differentially regulated (adjust p-value < 0.05 and |log2 fold change > 0.5|) between pig embryos and fetuses, where 447 genes were up-regulated in pig embryos and 1,258 genes were upregulated in pig fetuses (Figure 2).

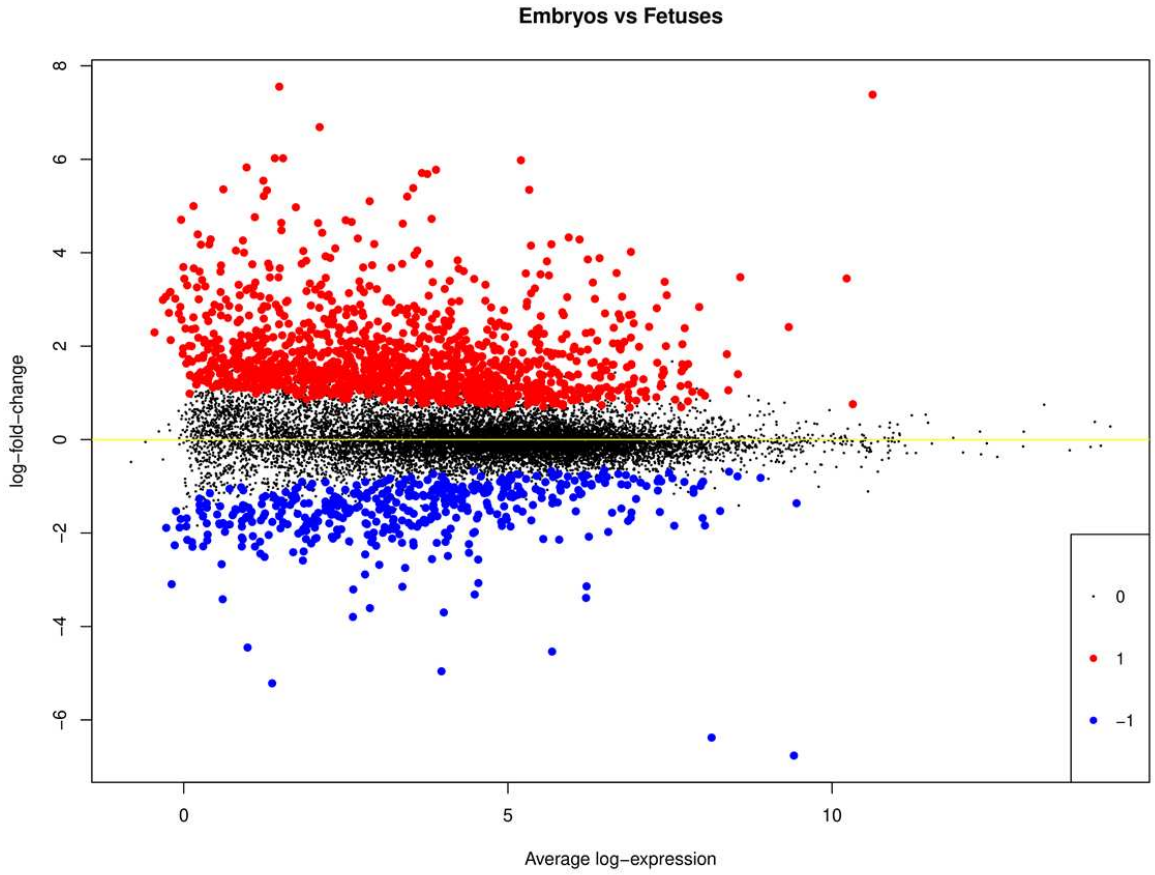


Figure 2: Mean-difference (MD) plot showing the log-fold change and average abundance of each differentially expressed genes between pig fetuses and embryos. Significantly upregulated genes in fetuses are highlighted in red (1) and upregulated genes in embryos are blue (-1), respectively, and non-differentially expressed genes in black (0).

The top 10 genes up-regulated in pig embryos and top 10 genes upregulated in pig fetuses according to the logFC (Table 1), were related to erythrocyte differentiation, heart development, ions transport, skeletal muscle system, neuronal development, ossification and renal function. The complete list containing all upregulated genes in pig embryos and upregulated in pig fetuses is available in Supplementary file 1 (https://drive.google.com/file/d/1zIILTT23duA1W85jE_N5l_GtbnEwlbZS/view?usp=sharing).

Table 1. Top 10 up-regulated genes in pig embryos and top 10 upregulated genes in pig fetuses

Upregulated genes in pig embryos				
Ensembl ID	logFC[1]	Gene Name	Gene Description	adj.*P.Val*[2]
ENSSSCG00000014726	-6.764396328	HBE1	Hemoglobin subunit epsilon 1	1.38E-08
ENSSSCG00000007975	-6.379504386	HBZ	Hemoglobin subunit zeta	8.50E-10
ENSSSCG00000031865	-5.217417343	AQP8	Aquaporin 8	8.04E-06
ENSSSCG00000021902	-4.96041836	GABRP	Gamma-aminobutyric acid type A receptor subunit pi	3.70E-12
ENSSSCG00000040513	-4.537759202	AQP3	Aquaporin 3	1.18E-09
ENSSSCG00000006731	-4.451641793	VTCN1	V-set domain containing T cell activation inhibitor 1	7.08E-10
ENSSSCG00000031080	-3.795925653	HAND1	Heart and neural crest derivatives expressed 1	6.49E-07
ENSSSCG00000000418	-3.699759861	TAC3	Tachykinin precursor 3	1.53E-11
ENSSSCG00000030461	-3.607432588	HEPHL1	Hephaestin like 1	2.04E-09
ENSSSCG00000002432	-3.418006247	KCNK13	Potassium two pore domain channel subfamily K member 13	1.87E-08
Upregulated genes in pig fetuses				
Ensembl ID	logFC[1]	Gene Name	Gene Description	*adj.P.Val*[2]
ENSSSCG00000009219	7.553512845	IBSP	Integrin binding sialoprotein	3.58E-12
ENSSSCG00000014725	7.383041924	HBB	Hemoglobin, beta	1.01E-10
ENSSSCG00000040098	6.688192867		Uncharacterized	1.21E-11
ENSSSCG00000037430	6.022006128	COL6A6	Collagen type VI alpha 6 chain	3.09E-10
ENSSSCG00000039501	6.020672325	NEUROD6	Neuronal differentiation 6	4.32E-08
ENSSSCG00000008072	5.978854537	ASPN	Asporin	1.58E-14
ENSSSCG00000009955	5.824620103	CRYBB2	Crystallin beta A2	9.52E-08
ENSSSCG00000035520	5.774356588		Uncharacterized	0.01103
ENSSSCG00000040641	5.70621141	CRYBA1	Crystallin beta A1	9.23E-05
ENSSSCG00000031903	5.684525698	TNNT3	Troponin T3, fast skeletal type	2.33E-12

[1]Log foldchange; [2]Adjusted *p-value* after multiple correction tests to reduce type I error; significance threshold: adjusted *p-value* ≤ 0.05

In functional analyses, a total of 203 GO terms associated with BP categories were enriched (FDR < 0.05) on David tool and submitted to Revigo tool, resulting in 71 BP (Figure 3). The enriched GO terms highlighted several BP crucial for developmental biology, such as anatomical structure development, regulation of cell growth, negative regulation of development process and regulation of biologic quality. Additionally, other BP related to the physiology of embryonic to fetal transition, including ossification, extracellular matrix organization, skeletal muscle development, erythrocyte differentiation, neuronal system and cardiovascular system, were enriched by many of the top DE genes between pig embryos and fetuses. More detailed information about GO terms are available in Supplementary file 2 (https://drive.google.com/file/d/1zIILTT23duA1W85jE_N5l_GtbnEwlbZS/view?usp=sharing).

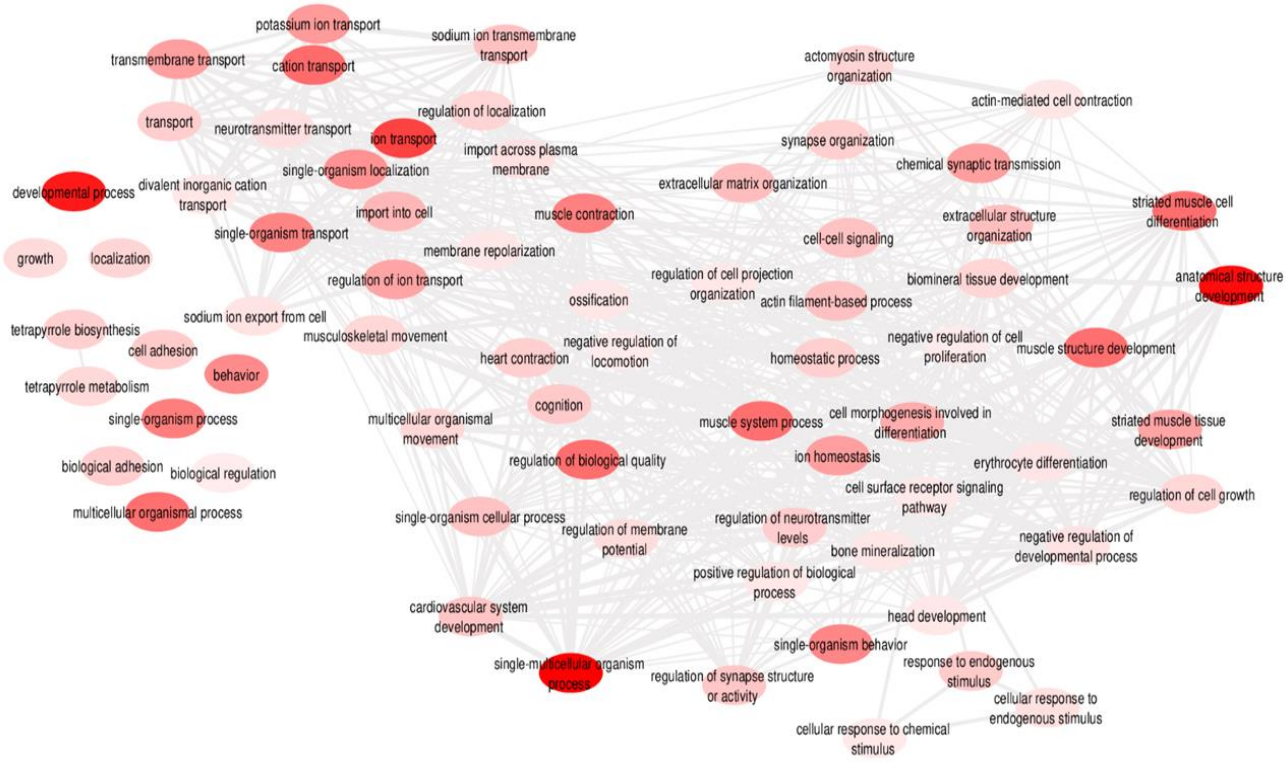


Figure 3: Biological processes (BP) (bubbles) network enriched for all differentially expressed (DE) genes between pig fetuses and embryos. Bubbles color indicates the adjust p-value of the gene ontology (GO) terms and the line width indicates the degree of similarity among GO terms. The design of this network was modified in Cytoscape version 3.8.0.

From the 173 pathways obtained in Toppcluster tool, 44 specific pathways were enriched by upregulated genes in pig embryos, among them are Hippo and TGF beta signaling pathways, pluripotency and proliferation pathways and cell-cell interaction. Even more pathways including myogenesis, collagen synthesis, and neuronal systems pathways (Figure 4) were enriched by upregulated genes in pig fetuses, which is line with the more differentiated tissues established at fetal development compared to embryonic stage. In addition, it was identified three pathways shared between upregulated genes in embryos and fetuses, related to transport of small molecules, transport of anions/cations and protein digestion and

absorption. Particularly, these pathways were enriched by genes that encoding different members of solute carrier (SLC) families, including the most upregulated gene in pig fetuses, the *COL6A6*. Additional details about the pathways enriched for embryos and fetuses' development and those that were shared between them are available in Supplementary file 3 (https://drive.google.com/file/d/1zIILTT23duA1W85jE_N5l_GtbnEwlbZS/view?usp=sharing).

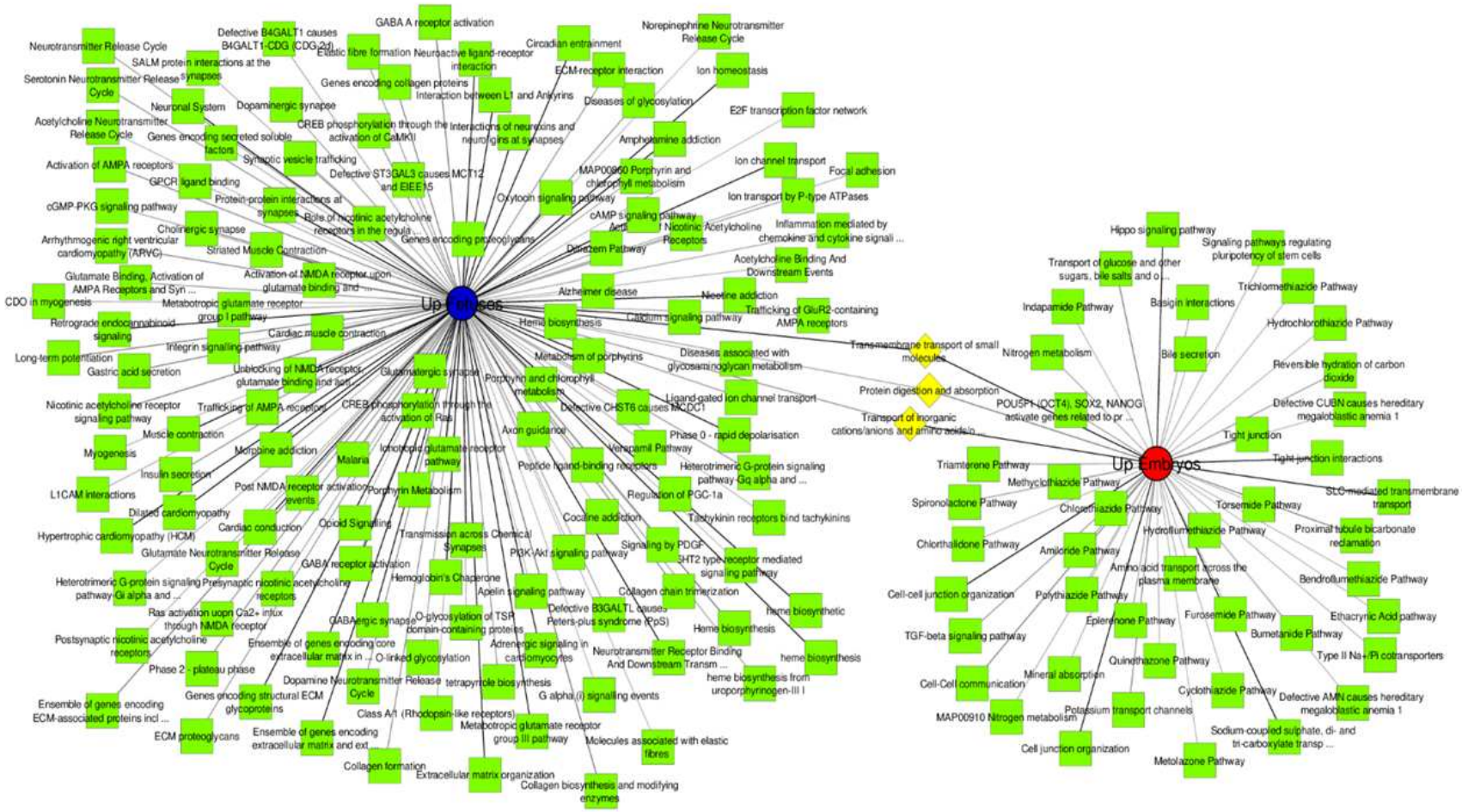


Figure 4: Pathway network for up-regulated genes in pig embryos and fetuses. All significant differential genes (HGNC Gene IDs) were classified as upregulated in embryos (cluster 1 in red) or in fetuses (cluster 2 in blue) and submitted to the ToppCluster online tool. Pathway category was used in functional enrichment and Cytoscape version 3.8.0 was used to modify the network design. Nodes for specific pathways are green and for shared pathways between embryos and fetuses are yellow.

Discussion

Organogenesis is a crucial period of the prenatal development, in which organs, tissues and system are established [1,28]. During this period and after, tight transcriptional control of genes involved in cell-cell interaction, cell proliferation and cell differentiation is required to ensure the normal development and functionality of the organs [29]. By using RNA-seq, it was possible to identify several differentially regulated genes during pig prenatal development, which were associated with BP and pathways, such as “biological regulation of proliferation and differentiation”, “cell morphogenesis”, “growth” and “cell-cell signaling”. These events are determinants for success on establishment of organ, tissues and structures in pig and other related species, such as in humans.

General signaling pathways during prenatal development

The Hippo pathway, which was initially discovered during organs growing in *Drosophila*, is highly conserved in mammalian, and its core is composed by an integrated cascade of kinase proteins that cooperatively regulate the transcriptional program of proliferation, differentiation and apoptosis [29,30]. In the present study, the Hippo signaling pathway was enriched by upregulated genes in pig embryos, including *FERMD1*, *CRB2*, *LLGL2, PARD6B, MPP5, PATJ, RASSF6* and *AJUBA* genes. These genes encode upstream modulatory components of the Hippo core, including FERM domain (FERMD1 and FERMD6), crumbs (CRB1, CRB2 and CRB3), LGL/LLGL Scribble Cell Polarity Complex (LGL1 and LGL2), PAR family (PAR3-6), Membrane Palmitoylated Protein 5 (PALS1/MPP5), PALS1-Associated

Tight Junction Protein (PATJ), RASSF family components (RASSF1-6) and AJUBA LIM Protein (AJUBA) [29,31]. Although the mechanism and critical physiologic functions of these upstream proteins in mammalian Hippo signaling are not completely clear, some of them are also components of cell-cell junctions (CRB1-3, PAR6, PALS1 and AJUBA) and cytoskeleton interaction (AJUBA and RASSF1-6), suggesting that the Hippo signaling cascade is modulated during the development by multiple inputs from cell density and extracellular environment interaction [29]. In line with this, in the present study, the Hippo related genes, namely *CRB2*, *MMP5*, *PARD6B* and *PATJ*, also shared the pathways tight junction and cell junction enriched for embryos development in the present study. Since events related to cell morphogenesis and cell differentiation are mediated from cell-cell contact mainly at embryonic development [3], the upregulation of genes related to Hippo pathway in pig embryos reinforce the transcriptional regulation from cell-cell interaction during embryonic stage.

The extracellular inputs are important to modulate the phosphorylation status of the kinases in the core of Hippo pathway and, thus controlling the activity of the downstream effectors, such as YAP (Yes1 Associated Transcriptional Regulator) [29], which was also upregulated in pig embryos and associated with Hippo pathway. The Hippo downstream effectors bind to transcription factors in the nucleus and promote the expression of genes related to cell proliferation and differentiation [29]. During the development, the YAP that can also act independently of the Hippo core cascade, has been associated with organogenesis

of heart, liver, kidney and brain (for review, see [32]), and cell fate determination in preimplantation embryos [33,34].

Therefore, the upregulation of Hippo related genes in pig embryos highlights the involvement of this pathway in biological events related to organ development, since multiple inputs related to cell interaction are crucial for organ and tissue morphogenesis, differentiation and apoptosis during embryogenesis [3]. Moreover, considering some divergences regarding the mammalian Hippo signaling involvement in the control of organ size or cell cycle exit (for review, see [30]), and that many authors suggest the deletion of Hippo components in multiple organ and at different times during development [29,30], our findings related to Hippo signaling enrichment in pig embryos may open possibilities for studies that address the main role of this signaling during organs development, using pig as an accessible model.

Despite the key role that Hippo signaling plays in transcriptional control, a possible crosstalk among Hippo and other signaling pathways, such as the BMP (Bone Morphogenic Protein)/ TGFβ (Transforming Growth Factor beta) pathway, can be required to enhance the transcriptional responses during critical stages of developmental biology [30,31]. The TGFβ and BMP proteins belong to TGFβ signaling superfamily, which are involved in cell patterning, organogenesis, and tissue homeostasis [35–37]. Based on this and in line with a possible crosstalk between signaling pathways, the TGFβ signaling pathway was enriched by upregulated genes in pig embryos, where *TGFBR1* (Transforming Growth Factor

Beta Receptor type 1) and *BMP4* (Bone Morphogenetic Protein 4) genes were also associated with Hippo signaling pathway.

The *BMP4* gene is expressed in several organs during mouse development, including heart, brain, muscle, and limb buds development [4], and during craniofacial development in avian embryos [38]. In these stages of development, the *BMP4* gene has been reported to be co-expressed with BMP and activin membrane-bound inhibition (*BAMBI*) gene in several embryonic organisms models [36,38–40]. The *BAMBI* gene encodes a transmembrane protein (BAMBI) structurally similar to receptors type I of TGFβ family, except by the kinase activity required for signaling [35]. This decoy receptor dimerizes with I and II receptors of TGFβ family, inhibiting the transcriptional response mediated by TGFβs and BMPs, which may be important for developmental homeostasis [38,41]. Hence, the *BAMBI* was also associated with TGF*β* signaling pathway together with *TGFβR1* and *BMP4* genes. These upregulated genes in pig embryos also shared several BP, including development processes, cell surface receptor signaling, cellular response to stimulus, cell morphogenesis involved in differentiation, biological regulation, positive regulation of biological processes and anatomical structure. Taken together, our findings highlight important mediators of the TGF*β* superfamily signaling, which are relevant for developmental homeostasis during organogenesis mainly from the control of cell proliferation, growth, and differentiation.

As long as the prenatal development advances, the maintenance of appropriated cell and tissue functions require greater degree of cell specialization, which depends on strict control of cellular fate, including proliferation and

differentiation [42,43]. The E2F family of transcription factors and its regulators are well-known in this context (for review, see [43]). During mouse organogenesis, E2F2 and E2F4 cooperatively act to maintain the proliferative stages of epithelial cells in tissues, such as skin and intestine [42]. The proliferative activity control of E2F2 is mediated exclusively by the RB (Retinoblastoma) phosphorylation [43]. The RB phosphorylation status during the cell cycle progression depends on the CDK/cyclins complex activity (cyclins-kinases dependents; CDK/ cyclins; *CCN* gene) (for review, see [43–45]), which is regulated by two families of Cdk inhibitors: the INK4 and Cip/Kip [46]. In line with this, the E2F transcription factor network was enriched by upregulated genes in pig fetuses, including *E2F2*, *RB1, CCND3* (D3-type cyclin), *CCNE2* (E2- type cyclin) and cyclin regulators, namely *CDKN1A* (Cyclin Dependent Kinase Inhibitor 1A, encoding a component of Cip/Kip inhibitor family) and *CDKN2C* (Cyclin Dependent Kinase Inhibitor 2C, encoding a component of INK4 inhibitor family) genes. Furthermore, BP such as biological regulation, negative regulation of cell proliferation, striated muscle cell differentiation and erythrocyte differentiation were enriched by *RB1*, *CDKN2C* and *CCNE2* genes. Taken together, these findings support that there is a tight transcriptional regulation to ensure the correct decision on cell fate during fetal stage, since specialized organs and tissues are formed and this control is crucial to maintain the adequate tissue function.

Another level of transcriptional control during development may involve histone modifying enzymes, such as the histone acetyltransferases (HATs), which potentializes the accessibility of transcriptional machinery in the promotor region of

the target gene [47–49]. However, some HAT, such as the p300/CBP-associated factor that contain a HAT domain, encoded by the *PCAF/KAT2B* gene, have been associated with acetylation of non-histone proteins [49], such as E2F members, particularly the E2F1, increasing its ability of DNA binding, stabilization and activation potential [50]. Hence, the *KAT2B* gene that was upregulated in pig fetuses is also associated with E2F transcription factor network, and shares with other E2F related genes (*RB1* and *CDKN2C*) the biological regulation and negative regulation of cell proliferation BPs. Moreover, recently *KAT2B* gene has been related to cell growth and differentiation during craniofacial and cartilage development in zebrafish and mouse embryos, which appears to be mediated by acetylation activity of the encoded protein [51]. These findings can reinforce important levels of gene transcription control related to proliferation and differentiation mediated by E2F transcription factors, as well as its patterns, which is in line with the development and maintaining of specialized organs and systems in fetal rather than embryonic stage, as highlighted in further sections of this article.

Cardiovascular system

As long as the development advances, the conceptuses nutritional requirements are not met by placental diffusion, demanding the establishment of the circulatory system [3,52]. In this context, the heart morphogenesis from a primitive tube to four-chambers based-system is required to allow the pumping of blood based on routes instead of unidirectional flow [52,53]. Among the marks of the cardiac chambers is the *HAND1* (Heart And Neural Crest Derivatives Expressed 1) gene, which is exclusively expressed in left ventricle during heart

development and induces the expression of cardiac genes [3,54]. The chamber septation is visible at similar ages in mammalian embryogenesis, such as in equine embryos at 25 days-old [55], dog and cat embryos around 23-25 days-old [2]. In pig prenatal development, the *HAND1* gene was one the most upregulated in embryos, suggesting the heart development and septation at embryonic stage, which is important to provide satisfactory nutritional support to many organs and tissues that have been formed and will continue to grow at fetal stage.

During the cardiovascular system development, members of semaphorin genes family has been related to blood vessels formation [53]. In this context, the cardiovascular system BP was enriched by members of semaphorins proteins class 3, namely, *SEMA3E* (semaphorin 3E) gene that was upregulated in pig embryos, and *SEMA3C* (semaphorin 3C) gene, which was upregulated in fetuses. The *SEMA3E* gene is associated with vascular patterning (vasculogenesis), in which the blood vessels are originated from endothelial progenitor cells (EPCs) or angioblasts, while the *SEMA3C* gene has been associated with angiogenesis, defined as new blood vessels formation from preexistent vessels (for review, see [53]). The vessel stabilization, homeostasis and anastomosis are mediated by Notch signaling pathway, while the response to angiogenic stimulus, which destabilize the quiescent vasculature and activate the endothelial cells differentiation to form new vessels, are stimulated by angiopoietin-2 (ANGPT2) [56]. In line with this, the cardiovascular system was also enriched by *NRARP* (Notch-regulated ankyrin repeat protein) and *ANGPT2* genes, which were upregulated in pig embryos and fetuses, respectively. The upregulation of genes

related to vasculogenesis (*SEMA3E* and *NRARP*) in embryos and upregulation of genes related to angiogenesis (*SEMA3C* and *ANGPT2*) in fetuses, suggest a differential regulation of the blood vessels formation during prenatal development in pigs, since vasculogenesis is an event primarily related to embryonic development, responsible to form primitives vascular plexus, while in the angiogenesis all blood vessels originate from preexistent blood vessels [52].

The first cells identified as blood cells during prenatal development are erythrocytes, which undergo many stages of differentiation until became mature and definitive blood cells [57]. The primitive erythrocytes, produced in yolk sac, mainly express genes of the globin epsilon (*HBE1*), a member of β-like globin, and beta (*HBZ*) subunits, a member of α-like globin, known as embryonic globin genes, while the definitive erythroid cells are produced in the fetal liver and mainly express fetal globin genes of gamma subunit (*HBG*) and adult globin, such as beta subunit (*HBB*) [58,59]. According to tight transcriptional control of globin genes during the development, these genes that enriched BP related to blood cells, including erythrocyte differentiation and transport, were the most differentially regulated in prenatal pig transcriptome, being the *HBZ* and *HBE1* genes the most upregulated genes in embryos, while the *HBB* gene was one of the most upregulated gene in fetuses. The switch from primitive to definitive erythrocytes mediated by globin gene expression is compatible with the quickly growth of fetal stage, since prenatal developmental progresses, more oxygen needs to be extracted from maternal blood, which is supported by enucleated definitive erythrocytes [57].

Furthermore, other DE genes required for blood cell formation were also associated with erythrocytes differentiation BP, including the *EPO* gene, which synthetases the erythropoietin, a cytokine that mediate the early signals needed to maturation of the primitive erythroid cells (for review, see [60]) and definitive erythropoiesis in mouse embryos [61], and thus was upregulated in pig embryos; and *GATA1* and *KLF1* genes, which coding the transcription factors GATA1 (GATA Binding Protein 1) and KLF1 (Kruppel Like Factor 1) involved on transcription of genes related to definitive erythropoiesis (for reviews, see [60,62]), and therefore, these coding genes were upregulated in pig fetuses.

As the circulatory system develops, which involves the establishment of heart, vessels and blood cell [3], as highlighted by the DE genes in this study, the heart presents fast and irregular contractions that start around 21-22 days of gestation in mammalian species [3,63,64]. However, in humans, the first expressive increase in the heart contraction is detected from embryos to fetuses transition, as a consequence of increased heart size and circulatory system maturation, which are events needed to ensure adequate blood supply for the body growth [63,65]. In line with the morphological changes and greater activity of the heart in fetal stage, the cardiac conduction pathway, were specifically enriched by upregulated genes in pig fetuses, including those coding by families of proteins involved in contractile activity, such as ATPases (*ATP1A2*, *ATP1B2*, *ATP2A1*, *ATP2B4*), Calcium Voltage-Gated Channel (*CACNA1D*, *CACNA1S*, *CACNA2D1*, *CACNA2D2*, *CACNA2D3*, *CACNB1*, *CACNG1*, *CACNG5*), Ryanodine receptor 3 (*RYR3*), alpha actin cardiac isoform (*ACTC1*) and troponin cardiac isoforms

(*TNNC1*, *TNNI1*). Furthermore, some of these genes, including *TNNC1*, *TNNT2* and *ACTC1*, enriched the Hypertrophic cardiomyopathy (HCM) pathway, which is associated with defectives in the sarcomeric proteins, resulting in failures on heart contraction as primary defect, followed by cardiac cells hypertrophy as compensation mechanism to this stress [66]. The enrichment of pathways related to cardiac conduction during pig fetal development is according with the greater nutritional apport required for the faster body growth during fetal stage.

Kidney development

The mammalian kidney development involves transition of the three main structures, the pronephros, mesonephro (intermediary kidney) and metanephros (definitive kidney) [55,67]. During this transition, a temporal overlapping, mainly related to mesonephro and metanephros development, is observed in several mammalian species, including human and pigs [68]. Therefore, many key factors have been described in both these stages of kidney development, including the transcription factors, such as SALL1 (Spalt Like Transcription Factor 1; *SALL1*) and WT1 (Wilms tumour suppressor; *WT1*), and morphogens inductors, such as WNT9b (Wnt Family Member 9B; *WNT9B*) [69,70]. Particularly, the SALL1 and WT1 are required for mesonephric tubules development, renal vesicles formation and metanephros function [70,71], and WNT9B, besides to be related to development of the mesonephric kidney tubules and initial responses to morphogenic stimulus need to renal vesicle formation [69–71], is also crucial to initial induction of nephron development, the structural and filtering units of the kidney [70]. The encode *SALL1*, *WT1* and *WNT9B* genes related to mesonephro

and metanephric development were upregulated in pig embryos at 25 days-old. This result can be explained by common role of the encoding proteins in mesonephro and metanephros development, since in pigs the development of mesonephro is observed to 15 from 24 days of gestation, overlapping with metanephros development, which starts around 20 days of gestation [68]. Furthermore, in the functional analyses, *SALL1* and *WT1* genes, share the anatomical structure, biological regulation and cell morphogenesis BP, and the *WNT9b* gene was associated with the “regulating pluripotency of stem cells signaling” pathway enriched for embryonic stage, which can be events associated with kidney development.

The nephron, which is a common filtering unit among stages of kidney organogenesis, is composed mainly by glomerulus, proximal and distal ducts, during early stages of kidney development [72]. The glomerulus is the most complex filtering structure, and podocytes are the main specialized cells that maintaining its function [69,73]. There are many genes expressed by glomerular podocytes during kidney development, including the abovementioned *WT1* gene, and *NPHS1* and *NPHS2* genes, encoding nephrin and podocin, respectively [3,69,74], which were also upregulated in pig embryos in the current study. Moreover, other important genes related to the physiologic function of the proximal and distal ducts in the prenatal kidney of humans were also upregulated in pig embryos. Among these upregulated genes in pig embryos, the members of Solute Carrier family (SLC), including *SLC9A2* (Solute Carrier Family 9 Member A2) and *SLC39A2* (Solute Carrier Family 39 Member 2) genes, are expressed in the

collecting duct in prenatal human kidney development, while *SLC3A1* (Solute Carrier Family 3 Member 1) and *SLC34A3* (Solute Carrier Family 34 Member 3) genes, are expressed in proximal duct in prenatal human kidney development [74]. Furthermore, the members of aquaporins (*AQP3*) and carbonic anhydrase (*CA2*), which are involved in water reabsorption and acid-basic equilibrium, are also considered as marker of collecting and proximal ducts in human kidney prenatal [74].

The upregulation of genes related to kidney filtering physiological function, mainly associated with functional markers of glomeruli (*NPHS1* and *NPHS2* genes) and distal and proximal tubules (SLC genes members, *AQP3* and *CA2* genes) in pig embryos, which enriched pathways related to renal function, such as nitrogen metabolism, mineral absorption, and proximal tubule bicarbonate reclamation, suggests the development and the high activity of the kidneys. During pig, human and sheep embryonic development, the intermediary kidneys are considered organs highly actives [68,75], which are well-developed at 25 days-old in pig embryos [75]. Taken together, our findings may indicate the transcriptional control compatible with the development and functionality of mesonephros during pig embryogenesis rather than fetal stage. Moreover, these findings provide valuable information for further investigation about this intermediary structure, since only recently the prenatal stages of kidney development have been reported in pig [75] and many of the current knowledge at molecular level of kidney organogenesis is based on metanephros [71,76].

Skeletal muscle development

Myogenesis is the most well-described biological event during the prenatal development in livestock, mainly because the number of muscle fibers is determined in the fetal stage and it is closely related to muscle mass [77,78]. The myogenesis transcriptional core is composed by muscle regulatory factors (MRFs), namely myogenic factor 5 (*MYF5)*, myogenic differentiation 1 (*MYOD*/*MYOD1*)*,* which are important for myoblasts differentiation from myogenic precursor, and myogenic factor 6 (*MYF6*/*MRF4*) and myogenin (*MYOG*), which are marker of multinucleated myotubes from myoblasts fusion [77,79–81]. The myotubes also express myosin heavy chain (MYH) isoforms, including the embryonic (*MYH3*) and fetal/neonatal (*MYH8*) [79]. The formation of the primary myotubes characterizes the first myogenic wave [79], which in pigs, occurs around 35-55 days of gestation [82]. Accordingly, the MRFs genes and *MYH3* and *MYH8* genes were upregulated in pig fetuses at 35 days-old. Furthermore, the MRFs genes enriched the myogenesis pathway together with *ME2FC* gene, which were also upregulated in pig fetuses. The *ME2FC* gene encodes one member of MEF2 transcription factor family, which is involved in the development of striated, cardiac and smooth muscles, as well as in neuronal cells (for review, see [83]). Particularly, in the skeletal muscle, the *MEF2C* acts jointly with MRFs genes to enhance and stabilize the expression of genes muscle-specifics during prenatal myogenesis in mouse [81]. According to functional cooperation among MRFs genes and *MEF2C* gene, many enriched BP were shared, including developmental process, growth, anatomical structure, biological regulation, striated muscle cell differentiation and striated muscle tissue development.

In order to provide a normal muscle development, the balance between positive and negative stimulus for cell proliferation and differentiation is required during prenatal development. In this context, the myostatin (*MSTN*) acts as a negative regulator of myoblasts proliferation, stimulating the muscle fiber differentiation [84] and the Myomaker Myoblast Fusion Factor (*MYMK/TMEM8C*) is a key muscle-specific factor required for myoblast fusion [85]. In line with the pig developmental stage in which the myoblasts proliferate and fuse, the *MSTN* and *MYMK* genes were upregulated in pig fetuses. Moreover, regarding the physiological role during muscle development and differentiation, *MSTN* and *MYMK* genes shared many BP, including developmental process, homeostatic process, muscle structure development and anatomical structure. These findings highlight the transcriptional control required for normal skeletal muscle development, which are mainly important during fetal development, when the biological events related to muscle cell proliferation, differentiation and fusion are more expressive.

The molecular mechanism that culminates in muscle growth is also mediated by the mitogenic factors, such as IGFs (insulin-like growth factors) that binds to IGF1R (insulin-like growth factor 1 receptor) [86,87]. The availability of IGFs is regulated by IGFBPs (Insulin Growth factors Binding Proteins), which can result in increased or reduced IGF release for binding to its receptor [86]. In this context, several IGFBP coding genes were upregulated in pig fetuses, including the *IGBP4*, *IGFBP5*, *IGFBP6* and *IGFBP7* genes that enriched the regulation of cell growth BP. Particularly, during myogenesis, the greater levels of IGBP-5 are

needed for myoblasts differentiation, since the increase in the IGBP-5 levels increases the local concentration of the IGF-2 ligand, increasing its release to IGFR1 (for review, see [86]). In line with this, the *IGFBP-5* gene shared response to endogenous stimulus, growth and biological regulation BPs with MRFs genes. Based on several genes that belong to IGFBP family that were upregulated in pig fetuses, and the lack of DE genes related to IGFs ligands between embryos and fetuses, it is suggested that IGF binding proteins rather than IGF itself can act as major molecular mediators of growth during pig prenatal development.

Skeleton development

The skeleton development occurs simultaneously with myogenesis, given the complementary functional properties of the musculoskeletal system, such as locomotion, protection of internal organs and body stability [88]. The early skeleton patterning involves the body segmentation in the three major skeletons, namely craniofacial, axial and appendicular, which is mediated by several transcription factor and signaling molecules [89]. Among them, the *BARX1* (BARX Homeobox 1), *MSX1* (Msh Homeobox 1), *ALX1* (ALX Homeobox 1), *HOXA1* (Homeobox A1), *HOXA2* (Homeobox A2) and *HOAX13* (Homeobox A13) genes which encode transcription factors that belong to Hox genes family (for review, see [89]), and *SHH* (Sonic Hedgehog Signaling Molecule) gene were upregulated in pig embryos. According to the role in the skeletal patterning during prenatal development, some of these genes enriched head development (*SHH*, *MSX1*, *HOXA2* genes) and anatomical structure development (*SHH*, *HOXA1*, *HOXA2*, *BARX1* and *MSX1* genes) BP. Besides these genes, the *TWIST1*/*TWIST* (Twist Family BHLH

Transcription Factor 1) and *DLL3* (Delta Like Canonical Notch Ligand 3) genes, which have been also associated with skeletal patterning [89], were also upregulated in pig embryos. These findings are accordingly with the body patterning from transcription factors and signaling molecules, which induce the specific gene transcription mainly at embryonic stages.

The ossification process in vertebrates can involve the endochondral or the intramembranous ossification [90,91]. Regardless of the ossification model, the bone matrix is synthetized by osteoblasts that differentiate from the mesenchymal progenitor cells [89]. Among the pathways that mediate the osteoblasts differentiation, the WNT-dependent signaling involve the activation of the Frizzled (FZD) receptor and/or the lipo-protein receptor-related protein (LRP5/LRP6) by WNT molecules binding and stabilization of the β-catenin, which stimulates the expression of the pro-osteogenic factors, including RUNX2 (RUNX Family Transcription Factor 2) and its downstream factor, OSX/SP7 (Osterix/Sp7 Transcription Factor) [92,93]. In the current study, the *WNT9A/WNT14* (Wnt Family Member 9A/14), *LRP5* and *SP7* genes which are components of WTN-signaling required for osteoblasts differentiation [94,95], were upregulated in pig fetuses. Moreover, the *WNT9A* and *FZD9* genes shared the cell surface receptor and anatomical structure development BP. Therefore, these results may provide an important molecular mechanism involved in osteoblast differentiation in pig fetuses.

Upon differentiation, the osteoblasts start to express markers involved on extracellular matrix mineralization, which is required for bone ossification [93]. In line with this, several genes encoding proteins involved on extracellular matrix

mineralization, including osteopontin, as known as secreted phosphoprotein 1 (*SPP1*), bone sialoprotein (*IBSP/BSP*), osteomodulin (*OMD*), dentin matrix acidic phosphoprotein-1 (*DMP1*) and phosphate-regulating gene with homologies to endopeptidases on the X chromosome (*PHEX*) [93,96–100], were upregulated in pig fetuses. Additionally, these genes enriched the ossification, biomineral tissue development and bone mineralization BPs, which are associated with its function. Based in our findings, it was possible highlight initial cues that can be crucial to onset of the bone development and mineralization at early pig fetuses, which are important events related to the embryonic-fetal transition in pig [101].

Extracellular matrix and prenatal development

The extracellular matrix (ECM) is copiously synthetized by differentiated cells to maintain the mechanical and biomechanical features of the specialized tissue, besides to acts as a microenvironment important to cell-cell signaling [102,103]. The most abundant component of the ECM is collagen, which is crucial to give stability and support to organs, such as bone, skin, kidney and liver [102,103]. Due to this, many of the DE genes in the current study encode members of collagen family, including members of collagens type V (*COL5A3)*, VI (*COL6A2*, *COL6A3* and *COL6A6*) and VIII (*COL8A1* and *COL8A2*), XI (*COL11A2*) and XII (*COL12A*), which were upregulated in pig fetuses and enriched collagen biosynthesis pathway and many others related to ECM assembly pathways.

Among these, the members of collagen type VI, including the up-regulated *COL6A2*, *COL6A3* and *COL6A6* genes in pig fetuses, are components of the ECM and are crucial to maintain the skeletal muscle functional integrity [104].

Particularly, the *COL6A6* gene, that was one of the most DE gene in pig fetuses, is widely expressed in human fetal tissues, including kidney and skeletal muscle [105]. Since myogenesis was a biological process more evidenced from MRF genes upregulated in pig fetuses at 35 days-old instead of embryos, and kidney for embryonic development, the upregulation of *COL6A6* gene in fetuses can suggest its major contribution during skeletal muscle development. Supporting to this, the collagen type IV (*COL4A4*), the unique member of collagen families that was upregulated in pig embryos, is one of the main components in glomerular region [106,107]. Indeed, the *COL4A4* gene mutation results in progressive loss of kidney function, characterizing the Alport's syndrome [106]. In addition, the *COL6A6* and *COL4A4* genes enriched the Protein digestion and absorption pathway, which was shared between embryonic and fetal development of pig. This pathway can reflect the functional involvement of the *COL6A6* and *COL4A4* genes in muscle and renal tissues metabolism, respectively.

The cartilaginous tissue is recognized to synthesize and secrete abundantly ECM components, mainly glycoproteins and proteoglycans, in addition to collagen proteins [108]. Among the non-collagen proteins, Matrilin 1 (*Crtm*/*MATN1*) is a glycoproteic component, while the aggrecan (*Agc1*/*ACAN*), is a proteoglycan almost unique to cartilage, which both are synthesized and secreted by highly active chondroblasts during chondrocytes differentiation [109]. In the pig prenatal transcriptome, the *MATN1* and *ACAN* genes were upregulated in pig fetuses, which can be associated with increased activity of differentiated chondroblast in pig fetuses at 35 days-old. Similarly, equine embryos at the same of prenatal

development presented abundant number of chondrocytes and chondroblasts in the bud of limbs and in some developing vertebrae [55].

The greater differentiation and activity of chondrocytes at pig fetal development can be associated with abundant cartilaginous ECM synthesized and used as a stable template for endochondral ossification, since in this process the cartilage undergoes remodeling to allow the invasion of ossification front cells, including osteoblasts, bone marrow and blood vessels [91,109]. This remodeling is mediated by the matrix metalloproteinase (MMPs) and the tissue inhibitors of metalloproteinases (TIMPS) [102]. Among these, the MMP2, which is encoded by *MMP2* gene upregulated in pig fetuses, have major role in the control of angiogenesis during endochondral ossification [109], while the TIMP2, which is encoded by TIMP2 gene that was also upregulated in pig fetuses, is the major expressed inhibitors of metalloproteinases in bone development [109]. According to its angiogenic control during endochondral ossification, the *MMP2* gene was associated with cardiovascular system and ossification BP, while the biological events potentially involved in the modulation of its proteolytic activity, such as response to endogenous stimulus and biological regulation, were enriched by *TIMP2* gene. Taken together, these findings can highlight the function of these genes during bone formation in pig fetuses at 35 days-old of gestation, trough ECM remodeling and angiogenic control.

Neuronal development

The neurogenesis is a crucial biological event responsible to the differentiation of stem cells or progenitor cells in neurons [110,111]. The neural

basic helix-loop-helix (bHLH) are important transcription factors required for inducing the neurogenesis program and activating the neuronal differentiation effectors, which are temporally expressed in order to ensure the correct neuronal identity during prenatal development [112]. Regarding this, members of neural bHLH were differentially regulated in pig prenatal transcriptome, including the *NEUROG1* (Neurogenin 1) gene, which was upregulated in pig embryos, and the *NEUROD6* (Neuronal Differentiation 6) gene, which was one of the most upregulated gene in pig fetuses. This differential expression reflects the functional role of neurogenin during embryogenesis, which acts as a cell fate determination factor and is exclusively expressed by neuronal proliferative cells [110], while the *NEUROD6* gene is exclusively expressed by differentiated neurons [111]. Therefore, the upregulation of the *NEUROG1* gene in pig embryos suggest an initial transcriptional program required for neuronal differentiation during embryogenesis, which was also evidenced by its association with Signaling pathways regulating pluripotency of stem cells pathway, while the upregulation of the *NEUROD6* gene in pig fetuses highlight the presence of differentiated neuronal cells at fetal stage.

Once differentiated, early neurons migrate from ventricular zone to primitive cortical plate [113]. During neurodevelopment, the neuronal migration is a well-described biological event, being characterized as radial and tangential migrations [114]. Basically, in the radial migration, excitatory neurons migrate from proliferative zones to the surface of the cerebral cortex, while in the tangential migration, inhibitory neurons migrate from ganglionic eminence also toward cortical

region [114]. The receptors for excitatory, namely, glutamatergic (Glutamate Ionotropic Receptor AMPA type – GRIA and Glutamate Ionotropic Receptor Kainate type - GRIK), and inhibitor neurotransmitters GABAergic (Gamma-Aminobutyric Acid Type A Receptor - GABR), are involved in the migration of cortical neurons at early development [115,116]. In pig prenatal transcriptome, many genes encoding different subunits of the glutamatergic receptors, such as *GRIA1*, *GRIA2*, *GRIA3*, *GRIA4*, *GRIK2*, *GRIK3*, and GABAergic receptors, such as *GABRA3*, *GABRA4*, *GABRA5*, *GABRB1*, *GABRB2*, *GABRB3*, *GABRG2*, were upregulated in pig fetuses and enriched the neuronal system pathway. In addition to these receptors, the gene encoding Roundabout Guidance Receptor 3 (*ROBO3*), which has been also related to cortex development and neuronal migration [117], was also upregulated in pig fetuses. In humans, an intense neuronal migration is observed at fetal stage in order to create a population of cortical neurons [113]. Our results suggest important signals to neuronal migration in pig fetal development rather than embryonic stage, since events related to neurogenesis was also highlighted for fetal stage in the present study, beside to reinforce the similarities between pigs and human regarding brain development.

Disturbs during neuronal migration impair the suitable balance between excitatory and inhibitory neurons in the cortex, which mainly affect the synaptic transmission [115,118]. In human prenatal development, the synaptogenesis starts from the embryos to fetuses transition as transitory circuits, which are important to ensure normal brain development and function [113,118]. In line with this early synaptogenesis, several upregulated genes in pig fetuses were associated with

synapse organization, regulation of synapse structure, regulation of synapses activity and chemical synapses transmission BP. The genes associated with these BP include, in addition to GABAergic and glutamatergic receptors encoding genes, genes that encode molecules directly involved in the synaptogenesis, including Synaptotagmin 1 (*SYT1*), Synapsin I (*SYN1*), Synaptophysin (*SYP*), leucine rich repeat transmembrane proteins (*LRRTMs*) and neurexins (*NRXN1*) [116,118], indicating the occurrence of events related to synaptic transmission during pig fetal development. On the other hand, the first synapses and synaptogenic molecules, including Synaptotagmin 1, Synapsin I and Synaptophysin, are detected only in mouse postnatal life [120]. The earlier prenatal synaptogenesis is related to greater complexity of the neurodevelopment [119]. Previous studies have reported the relevance of pig use in neuroscience, since pigs present similarities with human nervous system development rather than rodents species [120,121]. Therefore, our findings besides to suggest the transcriptional control related to synaptogenesis as an event required for proper neurodevelopment at fetal stage, reinforce the similarities between pig and human fetuses regarding neuronal development, which can feasible further studies addressing human prenatal issues related to synaptic plasticity, such as learning and memory [116].

Conclusion

In summary, this study provides a comprehensive landscape about the transcriptional changes associated with embryonic and early fetal development, which is important for understanding the key events mediating the morphogenesis and organogenesis in pigs. By the embryonic stage, genes related to Hippo

signaling and TGFβ/BMP4 signaling pathways evidenced events related to transcriptional regulation from morphogens, which are important mainly at early of prenatal development, while at fetal stage, pathways involved in cell differentiation for organs and tissues establishment were enriched by genes related myogenesis, extracellular matrix synthesis, and neuronal development. In addition, biological processes related to cardiovascular development and ossification processes reinforced important aspects of the developmental transition from embryos to fetuses. Therefore, the current study has emphasized data that may be relevant to future issues in developmental biology, instead of ensure a complete discussion of all molecular aspects involved in prenatal development.

References

[1] Edwards MJ, Saunders RD, Shiota K. Effects of heat on embryos and foetuses. Int J Hyperth 2003; 19:295–324.
[2] Pieri NCG, Souza AF, Casals JB, Roballo KCS, Ambrósio CE, Martins DS. Comparative Development of Embryonic Age by Organogenesis in Domestic Dogs and Cats. Reprod Domest Anim 2015; 50:625–631.
[3] Hyttel P, Sinowatz F, Vejlsted M, Betteridge K. Essentials of domestic animal embriology. Elsevier; 2010.
[4] Danesh SM, Villasenor A, Chong D, Soukup C, Cleaver O. BMP and BMP receptor expression during murine organogenesis. Gene Expr Patterns 2009; 9:255–265.
[5] Wesolowski SR, Raney NE, Ernst CW. Developmental changes in the fetal pig transcriptome. Physiol Genomics 2004; 16:268–274.
[6] Foxcroft GR, Dixon WT, Novak S, Putman CT, Town SC, Vinsky MD. The biological basis for prenatal programming of postnatal performance in pigs. J Anim Sci 2006; 84 Suppl:105–112.
[7] Ji Y, Wu Z, Dai Z, Wang X, Li J, Wang B, Wu G. Fetal and neonatal programming of postnatal growth and feed efficiency in swine. J Anim Sci Biotechnol 2017:1–15.
[8] Zeng S, Bick J, Kradolfer D, Knubben J, Flöter VL, Bauersachs S, Ulbrich SE. Differential transcriptome dynamics during the onset of conceptus elongation and between female and male porcine embryos. BMC Genomics 2019; 20:1–19.
[9] Lin H, Wang H, Wang Y, Liu C, Wang C, Guo J. Transcriptomic Analysis of

the Porcine Endometrium during Embryo Implantation. Genes (Basel) 2015:1330–1346.
[10] Zeng S, Ulbrich SE, Bauersachs S. Spatial organization of endometrial gene expression at the onset of embryo attachment in pigs. BMC Genomics 2019; 20:895.
[11] Te Pas MFW, De Wit AAW, Priem J, Cagnazzo M, Davoli R, Russo V, Pool MH. Transcriptome expression profiles in prenatal pigs in relation to myogenesis. J Muscle Res Cell Motil 2005; 26:157–165.
[12] Sollero BP, Guimarães SEF, Rilington VD, Tempelman RJ, Raney NE, Steibel JP, Guimarães JD, Lopes PS, Lopes MS, Ernst CW. Transcriptional profiling during foetal skeletal muscle development of Piau and Yorkshire-Landrace cross-bred pigs. Anim Genet 2011; 42:600–612.
[13] Verma N, Rettenmeier AW, Schmitz-Spanke S. Recent advances in the use of Sus scrofa (pig) as a model system for proteomic studies. Proteomics 2011; 11:776–793.
[14] Garcia IS, Teixeira SA, Costa KA, Marques DBD, Rodrigues G de A, Costa TC, Guimarães JD, Otto PI, Saraiva A, Ibelli AMG, Cantão ME, Oliveira HC, et al. l- Arginine supplementation of gilts during early gestation modulates energy sensitive pathways in pig conceptuses. Mol Reprod Dev 2020:1–16.
[15] Costa KA, Saraiva A, Guimaraes JD, Marques DBD, Machado-neves M, Reis LMB, Alberto F, Veroneze R, Oliveira LF de, Garcia IS, Teixeira SA, Guimarães SEF. Dietary L-arginine supplementation during early gestation of gilts affects conceptuses development. Theriogenology 2019; 140:62–71.
[16] Teixeira SA, Ibelli AMG, Cant E, Oliveira HC De, Coutinho LL, Guimar SEF. Sex Determination Using RNA-Sequencing Analyses in Early Prenatal Pig Development 2019.
[17] Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 2014; 30:2114–2120.
[18] Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12:357–360.
[19] Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics 2015; 31:166–169.
[20] BAQCOM. Bioinformatics Analysis for Quality Control and Mapping; 2019. https://github.com/hanielcedraz/BAQCOM. Accessed 15 February 2019.
[21] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43:1–13.
[22] R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing 2018.
[23] Benjamini Y, Hochberg Y. Controllin the false discovery rate: a practical and powerful approch to multiple testing. J R Stat Soc 1995; 57:289–300.
[24] Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2009; 35:169–175.
[25] Supek F, Bošnjak M, Škunca N, Šmuc T. Revigo summarizes and visualizes

long lists of gene ontology terms. PLoS One 2011; 6.
[26] Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ. ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and networkbased dissection of biological systems. Nucleic Acids Res 2010; 38:96–102.
[27] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A Software Environment for Integrated Models. Genome Res 2003; 13:2498–2504.
[28] Hu B, Wu X, Guo H, Wen L, Tang F, Fan X, Hu Y, Mao Y, Dong J. Single-cell RNA-seq analysis unveils a prevalent epithelial/mesenchymal hybrid state during mouse organogenesis. Genome Biol 2018; 19:1–20.
[29] Yu FX, Guan KL. The Hippo pathway: Regulators and regulations. Genes Dev 2013; 27:355–371.
[30] Halder G, Johnson RL. Hippo signaling: Growth control and beyond. Development 2011; 138:9–22.
[31] Pan D. The hippo signaling pathway in development and cancer. Dev Cell 2010; 19:491–505.
[32] Fu V, Plouffe SW, Guan K, States U. The Hippo pathway in organ development, homeostasis, and regeneration. Curr Opin Cell Biol 2017:99–107.
[33] Lorthongpanich C, Issaragrisil S. Emerging Role of the Hippo Signaling Pathway in Position Sensing and Lineage Specification in Mammalian Preimplantation Embryos1. Biol Reprod 2015; 92:1–10.
[34] Sharma J, Madan P. Characterisation of the Hippo signalling pathway during bovine preimplantation embryo development. Reprod Fertil Dev 2019.
[35] Tramullas M, Lantero A, Díaz Á, Morchón N, Merino D, Villar A, Buscher D, Merino R, Hurlé JM, Izpisúa-Belmonte JC, Hurlé MA. BAMBI (bone morphogenetic protein and activin membrane-bound inhibitor) reveals the involvement of the transforming growth factor-β family in pain modulation. J Neurosci 2010; 30:1502–1511.
[36] Karaulanov E, Knöchel W, Niehrs C. Transcriptional regulation of BMP4 synexpression in transgenic Xenopus. EMBO J 2004; 23:844–856.
[37] Dituri F, Cossu C, Mancarella S, Giannelli G. The Interactivity between TGFβ and BMP Signaling in Organogenesis, Fibrosis, and Cancer. Cells 2019; 8:1130.
[38] Higashihori N, Song Y, Richman JM. Expression and regulation of the decoy bone morphogenetic protein receptor BAMBI in the developing avian face. Dev Dyn 2008; 237:1500–1508.
[39] Tsang M, Kim R, De Caestecker MP, Kudoh T, Roberts AB, Dawid IB. Zebrafish nma is involved in TGFβ family signaling. Genesis 2000; 28:47–57.
[40] Grotewold L, Plum M, Dildrop R, Peters T, Rüther U. Bambi is coexpressed with Bmp-4 during mouse embryogenesis. Mech Dev 2001; 100:327–330.
[41] Yan X, Lin Z, Chen F, Zhao X, Chen H, Ning Y, Chen YG. Human BAMBI cooperates with Smad7 to inhibit transforming growth factor-β signaling. J Biol Chem 2009; 284:30097–30104.
[42] Dagnino L, Fry CJ, Bartley SM, Farnham P, Gallie BL, Phillips RA.

Expression patterns of the E2F family of transcription factors during murine epithelial development. Cell Growth Differ 1997; 8:553–563.
[43] Sears RC, Nevins JR. Signaling networks that link cell proliferation and cell fate. J Biol Chem 2002; 277:11617–11620.
[44] Wolgemuth DJ, Manterola M, Vasileva A. Role of cyclins in controlling progression of mammalian spermatogenesis. Int J Dev Biol 2013.
[45] Satyanarayana A, Kaldis P. Mammalian cell-cycle regulation: Several cdks, numerous cyclins and diverse compensatory mechanisms. Oncogene 2009; 28:2925–2939.
[46] Sage J. The retinoblastoma tumor suppressor and stem cell biology. Genes Dev 2012; 26:1409–1420.
[47] Gökbuget D, Blelloch R. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. Dev 2019; 146.
[48] Marzio G, Wagener C, Gutierrez MI, Cartwright P, Helin K, Giacca M. E2F family members are differentially regulated by reversible acetylation. J Biol Chem 2000; 275:10887–10892.
[49] Nagy Z, Tora L. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. Oncogene 2007; 26:5341–5357.
[50] Martínez-Balbás MA, Bauer UM, Nielsen SJ, Brehm A, Kouzarides T. Regulation of E2F1 activity by acetylation. EMBO J 2000; 19:662–671.
[51] Sen R, Pezoa SA, Shull LC, Hernandez-Lagunas L, Niswander LA, Artinger KB. Kat2a and Kat2b acetyltransferase activity regulates craniofacial cartilage and bone differentiation in Zebrafish and mice. J Dev Biol 2018; 6.
[52] Dabbagh A, Conte AH, Lubin L. Congenital heart disease in pediatric and adult patients: Anesthetic and perioperative management. 2017.
[53] Epstein JA, Aghajanian H, Singh MK. Semaphorin signaling in cardiovascular development. Cell Metab 2015; 21:163–173.
[54] Hu W, Xin Y, Hu J, Sun Y, Zhao Y. Inhibitor of DNA binding in heart development and cardiovascular diseases. Cell Commun Signal 2019; 17:1–11.
[55] Franciolli ALR, Cordeiro BM, da Fonseca ET, Rodrigues MN, Sarmento CAP, Ambrosio CE, de Carvalho AF, Miglino MA, Silva LA. Characteristics of the equine embryo and fetus from days 15 to 107 of pregnancy. Theriogenology 2011; 76:819–832.
[56] Murakami M. Signaling required for blood vessel maintenance: Molecular basis and pathological manifestations. Int J Vasc Med 2012; 2012.
[57] Dzierzak E, Philipsen S. Erythropoiesis: Development and differentiation. Cold Spring Harb Perspect Med 2013; 3:1–16.
[58] Nandakumar SK, Ulirsch JC, Sankaran VG. Advances in understanding erythropoiesis: evolving perspectives. Br J Haematol 2016; 173:206–218.
[59] McGrath KE, Frame JM, Fromm GJ, Koniski AD, Kingsley PD, Little J, Bulger M, Palis J. A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. Blood 2011; 117:4600–4608.
[60] Nandakumar SK, Ulirsch JC, Sankaran VG. Advances in understanding erythropoiesis: Evolving perspectives. Br J Haematol 2016; 173:206–218.

[61] Lee R, Kertesz N, Joseph SB, Jegalian A, Wu H. Erythropoietin (Epo) and EpoR expression and 2 waves of erythropoiesis. Blood 2001; 98:1408–1415.
[62] Sankaran VG, Xu J, Orkin SH. Advances in the understanding of haemoglobin switching: Review. Br J Haematol 2010; 149:181–194.
[63] Valenti O, Prima FAF Di, Renda E, Faraci M, Hyseni E, Domenico R De, Monte S, Giorgio E. Fetal cardiac function during the first trimester of pregnancy. J Prenat Med 2011; 5:59–62.
[64] Tan CMJ, Lewandowski AJ. The Transitional Heart: From Early Embryonic and Fetal Development to Neonatal Life. Fetal Diagn Ther 2019.
[65] Leiva MC, Tolosa JE, Binotto CN, Weiner S, Huppert L, Denis AL, Huhta JC. Fetal cardiac development and hemodynamics in the first trimester. Ultrasound Obstet Gynecol 1999; 14:169–174.
[66] Teekakirikul P, Kelly MA, Rehm HL, Lakdawala NK, Funke BH. Inherited cardiomyopathies: Molecular genetics and clinical genetic testing in the postgenomic era. J Mol Diagnostics 2013; 15:158–170.
[67] Seely JC. A brief review of kidney development, maturation, developmental abnormalities, and drug toxicity: Juvenile animal relevancy. J Toxicol Pathol 2017; 30:125–133.
[68] Moritz KM, Wintour EM. Functional development of the meso- and metanephros. Pediatr Nephrol 1999; 13:171–178.
[69] Schedl A. Renal abnormalities and their developmental origin. Nat Rev Genet 2007; 8:791–802.
[70] Little MH, McMahon AP. Mammalian kidney development: Principles, progress, and projections. Cold Spring Harb Perspect Biol 2012; 4:3.
[71] Georgas KM, Chiu HS, Lesieur E, Rumballe BA, Little MH. Expression of metanephric nephron-patterning genes in differentiating mesonephric tubules. Dev Dyn 2011.
[72] Raciti D, Reggiani L, Geffers L, Jiang Q, Bacchion F, Subrizi AE, Clements D, Tindal C, Davidson DR, Kaissling B, Brändli AW. Organization of the pronephric kidney revealed by large-scale gene expression mapping. Genome Biol 2008; 9.
[73] Roselli S, Gribouval O, Boute N, Sich M, Benessy F, Attié T, Gubler MC, Antignac C. Podocin localizes in the kidney to the slit diaphragm area. Am J Pathol 2002; 160:131–139.
[74] Wang P, Chen Y, Yong J, Cui Y, Wang R, Wen L, Qiao J, Tang F. Dissecting the Global Dynamic Molecular Profiles of Human Fetal Kidney Development by Single-Cell RNA Sequencing. Cell Rep 2018; 24:3554-3567.e3.
[75] Karalova E, Semerjyan Z, Manukyan A, Panyan N, Karalyan Z, Tatoyan M. The embryonic development of the pig excretory system. Porcine Res 2018; 8:1–12.
[76] Kuure S, Vuolteenaho R, Vainio S. Kidney morphogenesis: Cellular and molecular regulation. Mech Dev 2000; 92:31–45.
[77] Rehfeldt C, Fiedler I, Dietl G, Ender K. Myogenesis and postnatal skeletal muscle cell growth as influenced by selection. Livest Prod Sci 2000; 66:177–188.
[78] Maltin CA, Delday MI, Sinclair KD, Steven J, Sneddon AA. Impact of

manipulations of myogenesis in utero on the performance of adult skeletal muscle. Reproduction 2001; 122:359–374.
[79] Chal J, Pourquié O. Making muscle: Skeletal myogenesis in vivo and in vitro. Dev 2017; 144:2104–2122.
[80] Chang KC. Key signalling factors and pathways in the molecular determination of skeletal muscle phenotype. Animal 2007; 1:681–98.
[81] Wang DZ, Renee Valdez M, McAnally J, Richardson J, Olson EN. The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. Development 2001; 128:4623–4633.
[82] Wigmore PM, Stickland NC. Muscle development in large and small pig fetuses. J Anat 1983; 137 (Pt 2):235–45.
[83] Olson EN, Perry M, Schulz RA. Regulation of Muscle Differentiation by the MEF2 Family of MADS Box Transcription Factors. Dev Biol 1995; 172:2–14.
[84] Tasleem Jan A, Ju Lee E, Ahmad S, Choi I. Meeting the meat: delineating the molecular machinery of muscle development. J Anim Sci Technol 2016:1–10.
[85] Millay DP, O'Rourke JR, Sutherland LB, Bezprozvannaya S, Shelton JM, Bassel-Duby R, Olson EN. Myomaker is a membrane activator of myoblast fusion and muscle formation. Nature 2013; 499:301–305.
[86] Allard JB, Duan C. IGF-binding proteins: Why do they exist and why are there so many? Front Endocrinol (Lausanne) 2018; 9:1–12.
[87] Wang J, Feng C, Liu T, Shi M, Wu G, Bazer FW. Physiological alterations associated with intrauterine growth restriction in fetal pigs: Causes and insights for nutritional optimization. Mol Reprod Dev 2017; 84:897–904.
[88] DiGirolamo DJ, Kiel DP, Esser KA. Bone and Skeletal Muscle: Neighbors With Close Ties. J Bone Miner Res 2013; 28:1509–1518.
[89] Olsen BR, Reginato AM, Wang W. BONE DEVELOPMENT. Annu Rev Cell Dev Biol 2000:191–220.
[90] Arnold MA, Kim Y, Czubryt MP, Phan D, McAnally J, Qi X, Shelton JM, Richardson JA, Bassel-Duby R, Olson EN. MEF2C Transcription Factor Controls Chondrocyte Hypertrophy and Bone Development. Dev Cell 2007; 12:377–389.
[91] Mackie EJ, Ahmed YA, Tatarczuch L, Chen KS, Mirams M. Endochondral ossification: How cartilage is converted into bone in the developing skeleton. Int J Biochem Cell Biol 2008; 40:46–62.
[92] Long F, Ornitz DM. Development of the endochondral skeleton. Cold Spring Harb Perspect Biol 2013; 5:1–20.
[93] Rutkovskiy A, Stensløkken K-O, Vaage IJ. Osteoblast Differentiation at a Glance. Med Sci Monit Basic Res 2016; 22:95–106.
[94] Kato M, Patel MS, Levasseur R, Lobov I, Chang BHJ, Glass DA, Hartmann C, Li L, Hwang TH, Brayton CF, Lang RA, Karsenty G, et al. Cbfa1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. J Cell Biol 2002; 157:303–314.
[95] Day TF, Guo X, Garrett-Beal L, Yang Y. Wnt/β-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation

during vertebrate skeletogenesis. Dev Cell 2005; 8:739–750.
[96] Engin F, Yao Z, Yang T, Zhou G, Bertin T, Jiang MM, Chen Y, Wang L, Zheng H, Sutton RE, Boyce BF, Lee B. Dimorphic effects of Notch signaling in bone homeostasis. Nat Med 2008; 14:299–305.
[97] Qin C, Brunn JC, Cook RG, Orkiszewski RS, Malone JP, Veis A, Butler WT. Evidence for the proteolytic processing of dentin matrix protein 1. J Biol Chem 2003; 278:34700–34708.
[98] Qin C, D'Souza R, Feng JQ. Dentin Matrix Protein 1 (DMP1): New and important roles for biomineralization and phosphate homeostasis. J Dent Res 2007; 86:1134–1141.
[99] McKee MD, Cole WG. Bone Matrix and Mineralization. Second Edi. Elsevier Inc.; 2012.
[100] Ninomiya K, Miyamoto T, Imai J ichi, Fujita N, Suzuki T, Iwasaki R, Yagi M, Watanabe S, Toyama Y, Suda T. Osteoclastic activity induces osteomodulin expression in osteoblasts. Biochem Biophys Res Commun 2007; 362:460–466.
[101] Van Der Lende T, Van Rens BTTM. Critical periods for foetal mortality in gilts identified by analysing the length distribution of mummified foetuses and frequency of non-fresh stillborn piglets. Anim Reprod Sci 2003; 75:141–150.
[102] Jabłońska-Trypuć A, Matejczyk M, Rosochacki S. Matrix metalloproteinases (MMPs), the main extracellular matrix (ECM) enzymes in collagen degradation, as a target for anticancer drugs. J Enzyme Inhib Med Chem 2016; 31:177–183.
[103] Gelse K, Pöschl E, Aigner T. Collagens - Structure, function, and biosynthesis. Adv Drug Deliv Rev 2003; 55:1531–1546.
[104] Sabatelli P, Gualandi F, Gara SK, Grumati P, Zamparelli A, Martoni E, Pellegrini C, Merlini L, Ferlini A, Bonaldo P, Maraldi NM, Paulsson M, et al. Expression of collagen VI α5 and α6 chains in human muscle and in Duchenne muscular dystrophy-related muscle fibrosis. Matrix Biol 2012; 31:187–196.
[105] Fitzgerald J, Rich C, Zhou FH, Hansen U. Three novel collagen VI chains, α4(VI), α5(VI), and α6(VI). J Biol Chem 2008; 283:20170–20180.
[106] Miner JH. Organogenesis of the kidney glomerulus: Focus on the glomerular basement membrane. Organogenesis 2011; 7:75–82.
[107] Steven D. Funk*, Lin* M-H, Miner JH. Alport Syndrome and Pierson Syndrome: Diseases of the Glomerular Basement Membrane. Physiol Behav 2018.
[108] Lefebvre V, Smits P. Transcriptional control of chondrocyte fate and differentiation. Birth Defects Res Part C - Embryo Today Rev 2005:200–212.
[109] Ortega N, Behonick DJ, Werb Z. Matrix remodeling during endochondral ossification. Trends Cell Biol 2004; 14:86–93.
[110] Kim EJ, Hori K, Wyckoff A, Dickel LK, Koundakjian EJ, Goodrich L V, Johnson JE. Spatiotemporal Fate Map of Neurogenin1 (Neurog1) Lineages in the Mouse Central Nervous System. J Comp Neurol 2011; 519:1355–1370.
[111] Lee JE. Basic helix-loop-helix genes in neural development. Curr Opin

Neurobiol 1997; 7:13–20.
[112] Martynoga B, Drechsel D, Guillemot F. Molecular control of neurogenesis: A view from the mammalian cerebral cortex. Cold Spring Harb Perspect Biol 2012; 4.
[113] Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. Neuron 2016; 89:248.
[114] Hwang HM, Ku RY, Hashimoto-Torii K. Prenatal Environment That Affects Neuronal Migration. Front Cell Dev Biol 2019; 7:1–11.
[115] Cuzon VC, Yeh PW, Cheng Q, Yeh HH. Ambient GABA promotes cortical entry of tangentially migrating cells derived from the medial ganglionic eminence. Cereb Cortex 2006; 16:1377–1388.
[116] Luján R, Shigemoto R, López-Bendito G. Glutamate and GABA receptor signalling in the developing brain. Neuroscience 2005; 130:567–580.
[117] Barber M, Di Meglio T, Andrews WD, Hernández-Miranda LR, Murakami F, Chédotal A, Parnavelas JG. The role of Robo3 in the development of cortical interneurons. Cereb Cortex 2009; 19.
[118] Farhy-Tselnicker I, Allen NJ. Astrocytes, neurons, synapses: A tripartite view on cortical circuit development. Neural Dev 2018; 13:1–12.
[119] Chen VS, Morrison JP, Southwell MF, Foley JF, Bolon B, Susan A. Elmore. Histology Atlas of the Developing Prenatal and Postnatal Mouse Central Nervous System, with Emphasis on Prenatal Days E7.5 to E18.5 Vivian. Toxicol Pathol 2017:705–744.
[120] Conrad MS, Dilger RN, Johnson RW. Brain growth of the domestic pig (Sus scrofa) from 2 to 24 weeks of age: A longitudinal MRI study. Dev Neurosci 2012; 34:291–298.
[121] Aubid NN, Liu Y, Vidal JMP, Hall VJ. Isolation and culture of porcine primary fetal progenitors and neurons from the developing dorsal telencephalon. J Vet Sci 2019; 20:1–13.

CHAPTER 5

General conclusions

In the present study, we have demonstrated the pig prenatal transcriptome analysis by using RNA-seq tool. As a high throughput tool, we assessed all genes on all chromosomes in the genome of the interest specie, which allowed us to identify genes linked to sexual chromosomes that presented discrepant reads counts across the embryos and fetuses (conceptuses) samples and thus were used to perform the conceptuses sex determination. In the Chapter 2, we proposed the sex identification of pig conceptuses (embryos at 25 days-old and fetuses at 35 days-old) from read counts per million (CPM) of 10 genes mapped on Y chromosome, namely *DDX3Y*, *KDM5D*, *ZFY*, *EIF2S3Y*, *EIF1AY*, *LOC110255320*, *LOC110257894*, *LOC396706*, *LOC100625207* and *LOC110255257*, using available conceptuses RNA-seq data, with no need for extra techniques. This is valuable considering that in the evaluated prenatal stages of pig development, the phenotypic sex was not divergent among the samples. Additionally, from the genes used for conceptuses sex determination, is possible reinforce that molecular sex determination is a mechanism conserved across mammalian species, since the role of these genes is also important during male development in humans, highlighting the applicability of pigs as an animal model to study sex determination during human prenatal development.

In the context of maternal nutrition, several studies have reported the use of RNA-seq analysis as a powerful tool to identify the genes and molecular

mechanism related to fetal program. In the Chapter 3, we identified important changes on global embryos transcriptome at 25 days-old from gilts supplemented with 1.0% L-arginine maternal, while the global transcriptome of pig fetuses at 35 days-old was not affected by maternal dietary 1.0% L-arginine supplementation. At 25 days-old, we identified the downregulation of mitochondrial genes related to oxidative phosphorylation, namely *ND1*, *ND2*, *CYTB*, *COX2* and *ATP8*, and protein synthesis (16S rRNA similar sequence, *mt-rRNA type*), and upregulation of detoxification gene (*CYP1A1*), suggesting a regulation of energy metabolism in response to greater supply of nutrients in embryos and reinforcing an adjustment of metabolism in order to ensure homeostasis during embryos' development. On the other hand, the maternal dietary L-arginine supplementation did not affect the global transcriptome of the pig fetuses at 35 days-old, providing support for further discussions regarding the optimal period of supplementation in order to understand its impacts on swine prenatal development.

Regarding the developmental biology, many studies have reported gene expression analyses, using pig as model. However, these analyses are based on restrict set of known genes. Here, in the Chapter 4, by using RNA-seq analysis, we provided a general view about main transcriptional changes between pig embryos and fetuses, emphasizing data that may be relevant to future issues in developmental biology. From the 1,705 differentially regulated genes, we highlighted important biological events related to systems development, including cardiovascular, neuronal and skeletal muscle, which were more evidenced during fetal stage. On the other hand, more general biological events, such as cell

proliferation, cell-cell signaling, cell morphogenesis, were highlighted during embryonic development. The knowledge about these molecular changes is valuable to related species studies, such as human developmental biology, as well as to improve pig production from use of marker genes and their signaling pathways to investigate prenatal programming in the molecular context.