Nulliparity affects the expression of a limited number of genes and pathways in Day 8 equine embryos


To cite this version:

E. Derisoud, L. Jouneau, Catherine Archilla, Y. Jaszczyszyn, R. Legendre, et al.. Nulliparity affects the expression of a limited number of genes and pathways in Day 8 equine embryos. 2024. hal-04446593

HAL Id: hal-04446593
https://hal.inrae.fr/hal-04446593
Preprint submitted on 31 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License
Nulliparity affects the expression of a limited number of genes and pathways

in Day 8 equine embryos

E. Derisoud\textsuperscript{1,2}, L. Jouneau\textsuperscript{1,2}, C. Archilla\textsuperscript{1,2}, Y. Jaszczyzyn\textsuperscript{3}, R. Legendre\textsuperscript{6}, N. Daniel\textsuperscript{1,2}, N. Peynot\textsuperscript{1,2}, M. Dahirel\textsuperscript{1,2}, J. Auclair-Ronzaud\textsuperscript{6}, V. Duranthon\textsuperscript{1,2}, P. Chavatte-Palmer\textsuperscript{1,2}

1 Université Paris-Saclay, UVSQ, INRAE, BREED, Jouy-en-Josas, France
2 Ecole Nationale Vétérinaire d’Alfort, BREED, Maisons-Alfort, France
3 Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (i2BC), Gif-sur-Yvette, France
4 Institut Pasteur—Bioinformatics and Biostatistics Hub—Department of Computational Biology, Paris, France
5 IFCE, Plateau technique de la Valade, Chamberet, France

Correspondance:
Emilie Derisoud
emilie.deri soud@gmail.com
Bâtiment 230 – UMR 1198 BREED
INRAE - Domaine de Vilvert
22  78350 JOUY EN JOSAS, FRANCE

23

24 Pascale Chavatte-Palmer

25 pascale.chavatte-palmer@inrae.fr

26 Bâtiment 230 – UMR 1198 BREED

27 INRAE - Domaine de Vilvert

28  78350 JOUY EN JOSAS, FRANCE

29
Abstract

Nulliparous mares produce lighter and smaller foals compared to mares having previously foaled, with effects observed at least until 4 months of age. The need for a first gestation priming for the uterus to reach its full capacity has been proposed to explain this observation. Embryo developmental defects could be hypothesized but effects of maternal parity on the embryo have only been described once, in old mares, thus combining effects of parity and old age. The aim of this study was to determine effects of mare parity on embryo gene expression. Day-8 post ovulation blastocysts were collected from young (5/6 years old) nulliparous (YN, N=6) or multiparous (YM, N=4) non-nursing Saddlebred mares, inseminated with the semen of one stallion. Pure (TE_part) or inner-cell-mass-enriched (ICMandTE) trophoblast were obtained by embryo bisection for RNA sequencing (paired end, non-oriented, Illumina, NextSeq500). Deconvolution was performed on the ICMandTE dataset. Differential expression, with embryo sex and diameter as cofactors and gene set enrichment analysis (GO BP, KEGG, REACTOME databases) were performed using a false discovery rate <0.05 cutoff. Only a few genes were altered (ICM: n=18; TE: n=6) but several gene sets were perturbed (ICM: n=62; TE: n=50) by maternal parity. In YM, only pathways related to transcription, RNA processing and vesicle transport functions were enriched in the ICM whereas only pathways related to RNA localization were enriched in TE. In YN, while only gene sets related to ribosomes and extracellular matrix were enriched in the ICM, functions related to energy and lipid metabolism, lipid transport and interleukin-1 signaling were enriched in the TE. In conclusion, several genes and pathways are affected in embryos collected from nulliparous mares, with different effects on TE and ICM. Embryo development is altered in nulliparous mares, which could partially explain the term phenotype. Whether differences in gene expression result/induce poor embryo-maternal communication remains to be determined.
54 Keywords:

55 Blastocyst; RNA sequencing; horse; mare; periconception; equine
1. Introduction

In mammalian species, including the horse, it is now well established that the periconceptional and gestational maternal environment affect intra and extra-uterine growth and offspring long-term health [1,2]. These observations fall within the context of the Developmental Origins of Health and Diseases (DOHaD).

In horses, maternal parity defined as the number of gestations that produced a viable fetus (live or stillborn foal), is one of the main factors affecting the foal intra-uterine development. Indeed, foals born to primiparous mares (mares that have not foaled before) are lighter and smaller at birth and remain smaller until 18 months and lighter until 4 month of age compared to controls born to multiparous dams [3–13]. Their insulin sensitivity is higher than that of foals born to multiparous mares, and these data suggest that the normal decrease in insulin sensitivity observed in relation with foal age is delayed [13]. Similarly, testicular maturation is also delayed in foals born to primiparous mares [13]. These alterations in morphology and physiology of foals born to primiparous dams seem to be related to poorer performances in show jumping or on the racecourse than those of subsequent foals born to the same mare [14,15].

For a long time, these differences in mares’ first born foals have been attributed to the need for a first gestation priming for the uterus to be able to reach its optimal size and vascularisation and fully support feto-placental developmental needs [16]. Indeed, primiparous mares produced lighter and less voluminous placentas than multiparous ones [8,12,13,16]. In horses, placentation is diffuse and the epitheliocorial placenta is in contact with the entire surface of the uterus [17,18]. Most feto-maternal exchanges occur through branched vascular structures that form interdigitations with the mare endometrium, called microcotyledons, that maximize nutrient exchanges by increasing feto-maternal contact surface [17–19]. Reduced placental volume and weight are associated with reduced foal development in first born foals and suggest that primiparity could be a form of intra-uterine growth restriction in horses.
The placenta derives from the equine embryo trophoblast. Its later efficacy is conditioned by proper implantation and development. Implantation takes place around 35-38 days post ovulation [20]. Prior to that, the equine embryo develops free in the uterus and depends on direct support of uterine secretions for its development. Impaired pre-implantation development in nulliparous mares could play a role in the reduced size of both term placenta and newborn foal. The few existing studies that consider maternal parity on fertility effects are controversial. While some found that parity did not affect fertility [21–27], others reported that mares that have never foaled have reduced embryo and fetal mortality compared to mares that previously foaled [27–33]. Confounding effects of maternal parity and age is probably the source of those discrepancies. Indeed, in a recent epidemiological study considering the effect of parity only in mares older than 10 years, there is a cumulative negative effect of nulliparity and aging on the rates of pregnancy at 14 days post-ovulation (ED and PCP, personal communication). Maternal age have been shown to affect oocyte and embryo developmental capacities (for review [34]) as well as gene expression in Day 8 embryos [35]. At the opposite, only one study considered the effect of maternal parity on preimplantation embryo and showed alterations of the expression of genes related to embryo development and exchanges with the environment were observed [36]. This study, however, only considered mares older than 10 years, in which uterine degenerative changes had probably occur. As maternal age affects embryo gene expression, it is important to consider maternal parity in young mares. At this time, there is no study considering the effect of parity on gene expression of embryos in young mares.

The aim of this study was to determine the effect of maternal nulliparity in young mares on embryo gene expression at the blastocyst stage. Young (5-6 years old) nulliparous and multiparous mares were inseminated with semen of the same stallion. Day-8 blastocysts were collected, measured and bisected to separate the pure trophoblast (TE_part) from the inner cell mass enriched hemi-embryo (ICMandTE). Gene expression was analyzed by RNA-seq in each compartment.
2. Materials and methods

2.1. Ethics

The experiment was performed at the experimental farm of IFCE (research agreement C1903602 valid until March 22, 2023). The protocol was approved by the local animal care and use committee ("Comité des Utilisateurs de la Station Expérimentale de Chamberet") and by the regional ethical committee ("Comité Régional d’Ethique pour l’Expérimentation Animale du Limousin", approved under N° C2EA - 33 in the National Registry of French Ethical Committees for animal experimentation) under protocol number APAFIS#14963-2018050316037888 v2. All experiments were performed in accordance with the European Union Directive 2010/63EU. The authors complied with the ARRIVE guidelines.

2.2. Embryo collection

Twenty-one non-nursing mares (mostly French Anglo-Arabian with some Selle Français) aged from 5 to 6 years old were included in this study. Mares were allocated to one of 2 groups according to their parity: nulliparous (YN, n = 10) and multiparous mares (YM, n = 11). Multiparous mares were defined as dams that had already foaled at least once while nulliparous mares were defined as mares that had never foaled before the experiment. During the experimental protocol, mares were managed in one herd in natural pastures 24h/day with free access to water with no nutritional supplementation but for salt blocks. The experiments took place from April 1st to May 3rd, 2019. All mares remained healthy during this period. During the experimentation, mare’s withers’ height and weight were measured. Characteristics of all mares and mares that produced an embryo are detailed in Table 1.

Mares were monitored as previously described [35]. Briefly, the mares’ estrous period was monitored routinely by ultrasound with a 5MHz trans-rectal transducer. During estrus, ovulation was
induced with a single injection of human chorionic gonadotropin (i.v.; 750 - 1500IU; Chorulon® 5000; MSD Santé animale, France) as soon as one ovarian follicle >35mm in diameter was observed, together with marked uterine edema. Ovulation usually takes place within 48h, with > 80% occurring 25 to 48h after injection [37,38]. At the same time, mares were inseminated once with fresh or fresh overnight cooled semen containing at least 1 billion motile spermatozoa from a single fertile stallion. Ovulation was confirmed within the next 48 hours by ultrasonography.

Embryos were collected by non-surgical uterine lavage using prewarmed (37°C) lactated Ringer’s solution (B.Braun, France) and EZ-Way Filter (IMV Technologies, France) 10 days after insemination, i.e., approximately 8 days post ovulation. Just after embryo collection, mares were treated with luprotiol an analogue of prostaglandin F2α (i.m; 7.5 mg; Prosolvin, Virbac, France). The aim of the embryo collection was to obtain 5 embryos/group with each embryo coming from a different mare. Therefore, some mares that failed to produce an embryo at their first attempt were bred again for a second attempt.

2.3. Embryo bisection and RNA extraction

Using a binocular magnifying glass, collected embryos were immediately photographed with a size standard to subsequently determine embryo diameter using ImageJ® software (version 1.52a; National Institutes of Health, Bethesda, MD, USA). Embryos were then washed 4 times in commercially available Embryo holding medium (IMV Technologies, France) at 34°C and bisected with a microscalpel under binocular magnifying glass to obtain a trophoblast (TE_part) and an inner cell mass enriched (ICMandTE) hemi-embryo. At this stage, the TE_part is composed of trophectoderm and endoderm whereas the ICM is composed of epiblast layered on the internal side by endoderm cells [39,40]. Immediately after bisection, RNA extraction of each hemi-embryo was started in extraction buffer (PicoPure RNA isolation kit, Applied Biosystems, France) for 30 min at
42°C prior to storage at -80°C. RNA was extracted later from each hemi-embryo using PicoPure RNA isolation kit (PicoPure RNA isolation kit, Applied Biosystems, France), which included a DNase treatment, following the manufacturer’s instructions. RNA quality and quantity were assessed with the 2100 Bioanalyzer system using RNA 6000 Pico kit (Agilent Technologies, France) according to the manufacturer’s instructions.

2.4. RNA sequencing

Five nanograms of total RNA were mixed with ERCC spike-in mix (Thermofisher Scientific, France) according to manufacturer’s recommendations. Messenger RNAs were reverse transcribed and amplified using the SMART-Seq V4 ultra low input RNA kit (Clontech, France) according to the manufacturer recommendations. Nine PCR cycles were performed for each hemi-embryo. cDNA quality was assessed on an Agilent Bioanalyzer 2100, using an Agilent High Sensitivity DNA Kit (Agilent Technologies, France). Libraries were prepared from 0.15 ng cDNA using the Nextera XT Illumina library preparation kit (Illumina, France). They were pooled in equimolar proportions and sequenced (Paired end 50-34 pb) on NextSeq500 instrument, using a NextSeq 500 High Output 75 cycles kit (Illumina, France). Demultiplexing was performed with bc12fastq2 version 2.2.18.12 (Illumina, France) and adapters were trimmed with Cutadapt version 1.15 [41]. Only reads longer than 10 pb were kept.

2.5. RNA mapping and counting

As previously described [35], alignment was performed using STAR version 2.6 [42] on previously modified Ensembl 99 EquCab3.0 assembly and annotation. Genes were then counted with FeatureCounts [43] from Subreads package version 1.6.1.
2.6. Data analysis

All statistical analyses were performed by comparing YN to YM (YM set as reference group) using R version 4.0.2 [44] on Rstudio software version 1.3.1056 [45].

Embryo were sexed using X Inactive Specific Transcript (XIST) expression as previously described [35]. Six embryos were determined as female (2 in the YN group and 4 in the YM group) while 5 were considered as male (4 in the YN group, and 1 in the YM group).

2.6.1. Embryo recovery and fertility rate, embryo diameter and total RNA content analysis

Embryo recovery rates (ERR) per mare and per ovulation were calculated as the number of attempts with at least one embryo collected/total number of attempts. Both were analyzed using the Exact Fisher test to determine if maternal parity influenced embryo recovery.

For total RNA content analyses, as embryos were bisected without strict equality for each hemi-embryo, a separate analysis of ICM and TE and TE_par quantities would not have been meaningful. Thus, ICM and TE and TE_par RNA quantities were summed up. RNA quantity and embryo diameter were analyzed using a linear model of nlme package version 3.1-148 [46] including maternal parity and embryo sex, followed by 1000 permutations using PermTest function from pgirmess package version 1.6.9 [47]. Variables were kept in the subsequent models when statistically significant differences were observed. Differences were considered as significant for p < 0.05.

2.6.2. Deconvolution of gene expression in ICM and TE using DeMixT

The deconvolution method has already been described in equine embryos [35]. Briefly, this method enables the estimation of the relative gene expression of TE and ICM cell types within the hemi-
embryo ICMandTE which is composed of both trophoblast and inner cell mass in unknown relative proportions. After filtering all genes with 3 non-null count values in at least one group (YN or YM) per hemi-embryo (ICMandTE or TE_part), removing genes with a null variance in TE_part and adding the value “1” to all count values in ICMandTE and TE_part datasets, deconvolution was performed using the DeMixT R package version 1.4.0 [48,49]. Output datasets were DeMixT_ICM_cells and DeMixT_TE_cells, corresponding to the deconvoluted gene expression in ICM cells and TE cells of ICMandTE, respectively.

At the end of deconvolution, a quality check was automatically performed by the DeMixT R package with the TE_part used as reference for DeMixT_TE_cells. Genes were automatically filtered out if the difference between average deconvoluted expression of reference cells in mixed samples and average expression of reference cells > 4.

Outputs of DeMixT_ICM_cells vs DeMixT_TE_cells, DeMixT_ICM_cells vs TE_part and ICMandTE vs TE_part were compared with Deseq2 version 1.28.1 [50] to confirm that the deconvolution was effective at separating gene expression. To check if deconvolution was efficient, as previously described [35], the expression of several genes proper to ICM and TE cells in equine embryos identified using literature search [51] was compared before and after deconvolution. Results of these analyses were represented through manually drawn Venn diagrams as well as principal component analysis graphics of individuals, using ggplot2 version 3.3.3 [52] and factoextra version 1.0.7 [53].

### 2.6.3. Maternal parity comparison for gene expression

All genes with an average expression <10 counts in both YN and YM per hemi-embryo (ICM or TE) were filtered out on the DeMixT_ICM_cells and TE_part datasets. Differential analyses were performed with Deseq2 version 1.28.1 [50] with the YM group as reference, without independent filtering. Genes were considered differentially expressed (DEG) for FDR <0.05 after Benjamini-Hochberg correction (also known as false discovery rate, FDR). As ovulation was checked only every
48h and because embryos growth is exponential in the uterus, embryo diameter was considered as a cofactor in the model as well as embryo sex.

Equine Ensembl IDs were converted into Human Ensembl IDs and Entrez Gene names using gorth function in gprofiler2 package version 0.1.9 [54]. Genes without Entrez Gene names using gprofiler2 were manually converted when Entrez Gene names were available, using Ensembl web search function [55]. GO molecular function and GO Biological process annotations of genes were obtained from Uniprot website.

2.6.4. Gene set enrichment analyses (GSEA)

After log transformation using RLOG function of DESeq2 version 1.28.1, gene set enrichment analyses (GSEA) were performed on expressed genes using GSEA software version 4.0.3 (Broad Institute, Inc., Massachusetts Institute of Technology, and Regents of the University of California) [56,57] to identify biological gene sets disturbed by maternal parity. Molecular Signatures Databases [58] version 7.1 (C2: KEGG: Kyoto Encyclopedia of Genes and Genomes; REACTOME, C5: BP: GO biological process) were used to identify most perturbed pathways. Pathways were considered significantly enriched for FDR < 0.05. When the normalized enrichment score (NES) was positive, the gene set was enriched in the YN group while when NES was negative, the gene set was enriched in the YM group.

If applicable, as previously described in equine embryos [35], enriched terms from GO BP, KEGG and REACTOME databases were represented using SUMER analysis from SUMER R package version 1.1.5 and using FDR q-values [59]. Results were represented with graphs modified using Cytoscape version 3.8.2 [60]. In these graphs, gene sets are represented by nodes and the gene set size is represented by the size of the node. Node shape represents the gene set database (GO BP, KEGG or REACTOME). Blue nodes represent gene sets enriched in YN (NES > 0) while green nodes represent gene sets enriched in YM (NES < 0). Edge width represents the level of connection between representative
gene sets (thinner edges represent the first clustering while thicker edges represent the second clustering of the affinity propagation algorithm).

As SUMER is not able to consider only genes that participate to enrichment in GSEA, pathways with genes in common are grouped together, although genes in common are not the ones that participate to the enrichment. To better understand groups, therefore, authors, first recovered genes that were enriched in each pathway of a common group of gene sets according to SUMER analysis. Then, they only considered genes in common between pathways in one group to better qualify the function that was altered by maternal parity.

3. Results

3.1. Embryo recovery rates, diameter, total RNA content and quality and progesterone concentrations

Altogether, 25 embryo collections were performed (13 in YN and 12 in YM, 4 mares being flushed twice) and 12 embryos were obtained (7 from 6 YN mares and 5 from 5 YM mares). One young nulliparous mare produced twin embryos.

Embryo recovery rate per mare was 46% and 42% in YN and YM, respectively and did not differ between groups (p = 1).

Altogether, 1 and 2 double ovulations were observed, respectively, in YN and YM. The embryo recovery rate per ovulation at the time of embryo collection was not different according to group (50% in YN and 36% in YM, p = 0.70).

All embryos were expanded blastocysts grade I or II according to the embryo classification of McKinnon and Squires [61]. For the twin collection, embryos diameters were 580µm and 591µm. As only one embryo per mare was required, the 580µm diameter was randomly chosen for further analysis. Altogether, only 6 YN and all 5 YM embryos collected were RNA sequenced. Embryo diameter ranged from 457µm to 2643µm, with no effect of group on embryo diameter (p = 0.18). In
embryos selected for RNA sequencing, there was no effect of embryo sex on its size ($p = 0.63$). RNA yield per embryo ranged from 12.0 ng to 2915.5 ng and was not related to parity ($p = 0.07$) nor embryo sex ($p = 0.77$).

The median RNA Integrity Number (RIN) was 9.6 (8.9 - 10 range). Between 39.7 and 69.5 million reads per sample were obtained after trimming. On average, 70.94% of the reads were mapped on the modified EquCab 3.0 using STAR and 66.45% were assigned to genes by featureCounts.

### 3.2. Deconvolution of gene expression to discriminate ICM and TE gene expression in ICMandTE hemi-embryos

After selecting genes with more than 3 non null count values in at least one group (YN or YM) per hemi-embryo (ICMandTE or TE_part), 16,901 genes were conserved for deconvolution. In addition, 67 genes were removed because their variance was null in the TE_part. For these genes, the mean count in ICMandTE samples was above 110 counts. The deconvolution quality of all gene was sufficient. Therefore, at the end of the deconvolution algorithm, 16,834 genes were available for differential analysis.

Before deconvolution, 681 genes were differentially expressed (FDR < 0.05) between the ICMandTE and the TE_part (Fig 1a). After deconvolution, the comparison between DeMixT_ICM_cells and DeMixT_TE_cells yielded 6,171 differentially expressed genes while the comparison DeMixT_ICM_cells vs TE_part yielded 5,262 differentially expressed genes, with 4713 genes in common with the previous comparison (70%). Moreover, 677 of the initially 681 differentially expressed genes before deconvolution were also identified as differentially expressed in both post-deconvolution analyses. Only in the comparison DeMixT_ICM_cells vs TE_part, 3 among the 4 remaining genes were identified. On the PCA graph of individuals, ICMandTE and TE_part were partly overlapping (Fig 1b). DeMixT_TE_cells and TE_part superposed well, suggesting that datasets before and after deconvolution have a similar global gene expression; whereas the DeMixT_ICM_cells group is clearly separated from others on Axis 1 (22.3% of variance), indicating that the deconvolution
effectively enabled the separation of gene expression of the two cell types in the mixed part (ICM and TE).

On the 12 genes previously identified by Iqbal et al. as more expressed in the ICM [51], one had to be removed before deconvolution because its variance in the TE was zero (ENSECAG00000010653, annotated as SRY-Box Transcription Factor 2, SOX2). On the 11 remaining genes, 4 were also more expressed in the ICM vs TE_part comparison (Table 2). After deconvolution (comparison DeMixT_ICM_cells vs TE_part), 10 out of 11 of these genes were effectively more expressed in the ICM. Iqbal et al. identified 7 genes that were more expressed in the TE. One of those genes was differentially expressed in the comparison ICM and TE vs TE_part, i.e., before deconvolution. After deconvolution, the expression of 3 of the 7 reported genes, different from the only gene identified before deconvolution, were increased in the TE_part compared to the DeMixT_ICM_cells.

All of these results validate the deconvolution procedure and justify the use of data from the DeMixT_ICM_cells file. In the following results, the TE_part was used as representative of TE and DeMixT_ICM_cells was used as representative of gene expression in the ICM.

### 3.3. Sample selection

One embryo (YM) was larger than 2000 µm while all other embryos were smaller than 1400 µm in diameter (Supplementary Figure 1). Embryo size has been shown to affect equine embryo gene expression [62]. Thus, the analysis was performed both with or without this large embryo to check if results were affected. All but one differential expressed genes identified with the largest embryo were also found differentially expressed without it (Supplementary Figure 2). Nevertheless, to limit size effect, the analyses described below are those without the largest embryo, where only 6 YN and 4 YM embryos were analyzed. The results of the differential analyses that were performed including the 2643 µm large YM embryo are shown in Supplementary Tables 1 and 2.

### 3.4. Differential gene expression in deconvoluted ICM cells
After retaining only genes with an average expression ≥ 10 counts in at least one maternal parity group and hemi-embryo, 14,418 genes were considered as expressed in the YN or YM embryos ICM cells. Only 18 genes were differentially expressed (12 downregulated and 6 upregulated in YN) (Fig. 2 and Supplementary table 3). Respectively, 11 and 5 genes out of the down- and upregulated genes were associated to a protein known and described in human. These 16 genes an gene sets determined from Uniprot in which they are susceptible to play a role are presented in Table 3.

3.5. Differential gene expression in the TE part

In the TE, 13,203 genes were considered as expressed in YN or YM. Only 6 were differentially expressed (Supplementary table 4) with half being down and up-regulated in YN (Fig. 2). Except one that was a long noncoding RNA, all other genes were associated to a known protein in human. These genes are presented in Table 4 with the pathways in which they are susceptible to play a role.

3.6. Gene set enrichment analysis in deconvoluted ICM cells

After Entrez Gene ID conversion, 12,892 genes were considered expressed in ICM cells. Fifty-eight GO Biological Process and 4 KEGG pathways were disturbed by maternal parity in ICM cells (Supplementary table 5). After SUMER analysis, 2 and 27 gene sets, respectively enriched in YN and YM, were represented (Fig. 3). They were clustered in 8 groups. The group enriched in YM and clustered under the term “DNA recombination” was composed of genes related to the maintenance of DNA integrity, chromosome segregation and recombination. Enriched in YM groups “NCRNA metabolic process” and “Peptidyl lysine trimethylation” contained both, genes related to methylation and transcription. The only gene set enriched in YN in the group “NCRNA metabolic process” was mainly enriched by genes encoding for a subunit of ribosomes that were common with other gene sets enriched in YM. Genes related to ribosomes were, however, not participating in gene set enrichment of other pathways in this cluster. “multi organism localization”, “RNA splicing” and “RNA localization” clusters were composed of genes involved in RNA maturation and transport. The last
group enriched in YM was clustered under the term “vesicle targeting” and was containing genes related to intracellular transport. In this group, the pathway “Golgi vesicle transport” included a DEG that was Vacuolar Protein Sorting-Associated Protein 52 Homolog (VPS52), up-regulated in the ICM of embryos from YM mares. The only one group enriched in YN was composed of one pathway named “ECM receptor interaction” in which genes related to extracellular matrix (ECM) were observed.

3.7. Gene set enrichment analysis in TE

After Entrez Gene ID conversion, 11,889 genes were considered expressed in TE from YN or YM embryos. Altogether, 50 gene sets from GO BP, KEGG and REACTOME were perturbed (23 GO BP, 7 KEGG and 20 REACTOME) by maternal parity in young mares (Supplementary table 6). After SUMER analysis, 36 gene sets were represented (Fig. 3) and were clustered in 8 groups. Among them, 7 were enriched in YN. The first group was composed of one gene set named “cardiac septum morphogenesis”. Most genes that participated to the enrichment of this pathway were related to transcriptional factors. The second group was composed of 3 pathways and clustered under the term “negative regulation of secretion”. Genes that participated most to the enrichment of these pathways were related to the innate immunity, more particularly to the production and transport of the interleukin 1 beta (IL1B). Altogether, 3 groups were related to the production of energy inside the cell. Indeed, the cluster under the term “Parkinsons disease” was actually composed of genes with an enriched expression that were related to oxidative phosphorylation. One group was “Hydrogen peroxide metabolic process” where genes that participated the most to the enrichment were involved in the degradation of hydrogen peroxide. The last group involved in energy production was clustered under the term “valine, leucine and isoleucine degradation”. Genes that participated the most to the enrichment of these pathways were directly involved in the beta oxidation of fatty acid. The group under the term “positive regulation of lipid transport” was composed of genes related to the regulation of the transport of lipids and cholesterol. The last group enriched in YN was
clustered under the term “eukaryotic translation elongation” and was mostly enriched due to components of the ribosomes that mostly participate in these enrichments. Moreover, the REACTOME pathway “RRNA modification in the nucleus and cytosol” was clustered with this group because of genes that encoded for ribosome components. These genes were, however, not enriched in this particular pathway. Genes that participate to its enrichment, nevertheless, were related to ribosome biogenesis. The only one group enriched in YM was represented by the term “Regulation of glucokinase by glucokinase regulatory protein”. These pathways were mostly enriched in YM because of genes that encode for nucleoporin subunits.

4. Discussion

Maternal parity in young mares slightly affected both ICM and TE gene expression without affecting embryo recovery rates nor growth. Although only a few genes were affected by maternal parity, up regulated genes in the ICM of embryos from young nulliparous were involved in lipid, amine and creatinine metabolism, positive regulation of transcription, growth factor signaling and morphogenesis while down regulated were related to extracellular matrix (ECM) disassembly, reactive oxygen species (ROS) metabolism, transcription regulation, endocytosis, protein transport, protein metabolism, MAP kinase signalization and cell differentiation. In the TE, only five known genes were observed differentially expressed. One of the 3 up regulated genes in the TE of embryos from YN was involved in the cell response to hypoxia while the 2 others encode for ion binding proteins. Down regulated genes in the TE of embryos from YN mares were related to prostaglandin metabolism and amino acid/creatine exchange. Interestingly, while gene set enrichment analysis showed almost only pathway enrichment in YM in the ICM, in the TE, almost all pathways were enriched in YN embryos. In the ICM, after SUMER analysis, the pathways enriched in YM were related to DNA modification, RNA production and maturation and cell transport while gene sets related to the extracellular matrix function and ribosome were enriched in ICM of YN embryos. In the TE, gene
sets enriched in embryos from YN were related to immunity, growth factor signaling, phosphorylation oxidative, metabolism of reactive oxygen species, beta oxidation and transport of lipids and ribosome while gene sets enriched in embryos from YM were related to nucleoporins.

In the ICM of embryos from YM mares, enriched gene sets were mostly related to DNA conformation and methylation changes as well as RNA formation, transport and maturation. These results suggest that transcription and regulating pathways are less active in embryos from nulliparous mares. These results are comforted by the fact that teneurin transmembrane protein 3 (TENM3) is downregulated in embryos from nulliparous mares. This gene is part of the teneurin family, which encodes for transmembrane proteins that are essential for embryo morphogenesis and nervous system development. The knockdown of these genes in mice and drosophila leads to embryo lethality (for review [63]). Altogether, these results suggested that ICM growth and development would be poorer in embryo from nulliparous compared to multiparous mares but, here, no difference in embryo size had been observed. One hypothesis could be because only ICM seemed affected or because, embryo size at a same age is highly variable as shown in several studies [64–68], although ovulation check was performed twice daily [69]. This huge variation could hide size differences in studies.

In the TE, the gene named “family with sequence similarity 162 member A” (FAM162A), also known as E2-Induced Gene 5 Protein (E2IG5) or growth and transformation-dependent protein (HGTDP) was up-regulated in embryos from nulliparous mares. This gene is one of the hypoxia inducible factors (HIF)-activated downstream gene and is normally responsible of the activation of mitochondrial proapoptotic cascades when overexpressed [70]. As energy production processes (protein and lipid oxidation, oxidative phosphorylation and related regulatory pathways) were enriched in embryos from nulliparous mares, it seemed unlikely that FAM162A up-regulation in embryos from nulliparous mares was a response to hypoxic environment. Nevertheless, it could be hypothesized that the uterine environment of embryos may vary according to mares’ parity, partly due to reduced uterine blood perfusion in nulliparous mares.
study on the effect of nulliparity on uterine vascularization in young mares. It has nonetheless been
shown that FAM162A expression is increased in intestinal and uterine cervical cancer [71,72] its
overexpression enhanced cell proliferation processes, suggesting a non-elucidated positive role in
tumor development [72]. As in tumor, here, FAM162A could play a role in cell proliferation of equine
embryos but the process remains to be elucidated. This could also indicate that proliferation in
equine embryos differ according to mares’ parity, maybe as a response to their environment.

In the TE of embryos from nulliparous mares, pathways related to oxidative phosphorylation were
enriched. The enrichment of the expression of genes involved in these pathways could indicate that
the production of ATP from oxidative phosphorylation is up regulated in TE of embryos from
nulliparous mares in comparison to the ones from multiparous mares. This up-regulation of oxygen
oxidation in mitochondria could be harmful for TE cells as oxidative phosphorylation is accompanied
by the production of reactive oxygen species (ROS) and particularly of hydrogen peroxide (for review
[75]). Pathways related to hydrogen peroxide metabolic processes, however, were also enriched in
the TE of embryos from nulliparous mares, showing that there is an up-regulation of the control of
ROS such as hydrogen peroxide. The up regulation of both oxidative phosphorylation and regulation
of ROS pathways suggests that there is an increased production of energy that is not harmful
because well controlled in the TE of embryos from nulliparous mares.

At this developmental stage in equine embryos, 40 to 50% of glucose uptake is oxidized in the
mitochondria, probably to meet the high energy demand of ionic transport associated with the
important growth of both blastocoelic cavity and trophoblast [76]. Here, however, the enrichment in
oxidative phosphorylation was not accompanied by an enrichment in glucose metabolism nor
transport pathways but pathways linked to beta-oxidation of lipids and degradation of amino acids
were enriched. This suggests that glycolysis is not affected by maternal parity but, to meet energy
requirements, embryos from nulliparous mares use fatty acids and/or amino acids whilst embryos
from multiparous mares do not need more energy than already provided and therefore, do not
require the degradation of these substrates. The increased catabolism of amino acids and lipids could
be detrimental for embryo development as the first are required for protein synthesis and the latter are mandatory for hormone production (for review [77,78]). Pathways related to amino acid degradation, however, were mostly enriched in genes involved in beta-oxidation of fatty acids. Moreover, pathways related to translation and protein maturation but not pathways related to amino acids transport were enriched in embryos from nulliparous mares. Altogether, these results suggest that only lipid catabolism is enriched in the TE of nulliparous mares’ embryos.

In addition, pathways related to the transport of lipids and cholesterol were enriched in the TE of nulliparous mares’ embryos compared to those of multiparous mares. One hypothesis to explain these results could be that there is a higher energy demand in embryos from nulliparous mares and that they would compensate by degrading more lipids for oxidative phosphorylation, which requires more lipids to be obtained from the external environment. Another possibility is that the lipid composition of the uterine environment is altered in nulliparous mares, possibly due to immature uterine glands, leading to increased absorption by the embryo, that would stimulate beta-oxidation and thus oxidative phosphorylation. Indeed, the metabolism of lipids was also shown to be perturbed in blastocysts at the same developmental stage according to maternal parity in old mares [36]. To the authors’ knowledge, there is no study on the effects of maternal parity in any species on uterine fluid composition and how it could interfere with embryo gene expression. Although it is more likely that there are modifications in the uterine environment according to mare’s parity, the present results cannot conclude about the origin of the altered embryo metabolism.

As a confirmation of increased lipid transport, retinol binding protein 1 (RBP1) was up regulated in the ICM of embryos from nulliparous mares. Retinol is well known to be an important regulator of vertebrate development (for review [79]). In bovine, the addition of retinol to the maturation and culture medium of oocytes and embryos increased the blastocyst rate [80]. RBP transports the hydrophobic retinol in physiological fluids such as plasma [81] or uterine fluids. Pig conceptuses at the time of elongation produce RBP in large amounts, suggesting that retinol is important for embryo development [82]. In horses, the expression of RBP increased in the endometrium during diestrus.
under steroid regulation but did not vary according to the presence of an embryo or not [83].

Although underlying mechanisms are missing, RBP1 could, however, play an important role in equine early embryo development by transporting retinol to the embryo. The increased expression of RBP in the ICM of embryos from nulliparous mares could be a response to a reduced availability of retinol in the close environment of the embryo or an increased requirements of retinol from embryos of nulliparous mares.

Nutrient and ion exchanges were also modified by maternal parity in the TE. Indeed, solute carrier family 47 member 1 (SLC47A1 also known as MATE1), the solute transporter for molecules such as creatinine or guanidine, was down regulated in the TE of embryos from nulliparous mares. In addition, the expression of EF-hand calcium binding domain 11 (EFCAB11) and GTP binding protein 8 (GTPBP8) was increased in the TE of embryos from nulliparous mares compared to that of multiparous mares. These results could indicate a perturbed transport of different molecules in the TE of embryos from nulliparous mares.

These modifications of cell metabolism in the TE were associated with an alteration of pathways related to immunity, especially those linked to interleukin 1 beta (IL1B), being enriched in embryos from nulliparous mares. In cattle, it has been suggested that the early bovine embryo interacts with the dam’s immune system through processes involving IL1 [84]. In horses, maternal recognition of pregnancy (MRP) is thought to take place between 10-13 days post ovulation (for review [85]). At 19- and 25-days, but not at 13 days post ovulation, expression of the IL1 receptor antagonist has been shown to be markedly increased in the endometrium of pregnant compared to cyclic mares, suggesting that the endometrium regulates the IL1 signal and that IL1 plays a role in MRP in equine [86]. The expression of IL1B is increased in the luminal epithelium of pregnant vs cyclic mares at 10-13 days post ovulation, confirming the involvement of IL1B signaling process in MRP [87]. Here, embryos were collected earlier from the assumed MRP period but the observed differences in the IL1B signaling pathway could indicate that embryo-maternal communication and possibly MRP are affected by maternal parity.
Furthermore, related to lipid metabolism and IL1B signaling, peroxiredoxin like 2B (PRXL2B), also known as Prostamide/Prostaglandin F Synthase, was downregulated in the TE of nulliparous mares’ embryos. This gene encodes for an enzyme that has been shown to catalyze the reduction of prostamide $H_2$ to prostamide $F_{2\alpha}$ as well as the reduction of prostaglandin $H_2$ (PGH$_2$) to prostaglandin $F_{2\alpha}$ (PGF$_{2\alpha}$) [88]. In bovine, IL1B upregulates PGF$_{2\alpha}$ and prostamide secretion by in vitro cultured endometrial cells [89]. In horses, PGF$_{2\alpha}$ is secreted by the uterus to provoke the corpus luteum luteolysis (for review [90]). It has been shown that the suppression of the pulsatile secretion of PGF$_{2\alpha}$ from the endometrium is responsible for the maintenance of pregnancy [91] and that in vitro, PGF$_{2\alpha}$ production is significantly reduced when endometrial explants are co-cultured with embryonic tissues [92]. From the oviduct stages, equine embryos are able to produce prostaglandins [93–95]. Prostaglandins produced by the embryo, however, do not reach the blood circulation in sufficient amount to induce luteolysis [91]. It has been shown that these prostaglandins are required for myometrial contractions that participate in the migration of the equine embryo at the time of MRP [96]. By impeding the movement of the embryo, one study observed that equine embryo migration through at least 2/3 of the uterus is required to prevent luteolysis [97]. Moreover, the use of an intra-uterine device to imitate the physical presence of an embryo, allowed to prevent the luteolysis [98]. A recent study, nevertheless, observed that the contact of a substance/object is not sufficient to reduce PGF secretion from the endometrium, suggesting that embryo secretions are required for luteolysis [99]. Therefore, although MRP is thought to begin 2 days later, the present study shows that MRP might be delayed or disturbed in nulliparous mares.

5. Conclusion

So far, the effect of mare’s parity on embryo gene expression had never been considered. The present study shows that mare’s parity affects the expression of genes in both ICM and TE of blastocysts. Only the expression of few genes is altered but several important functions for embryo
development are affected by mare’s parity. Indeed, nulliparity in young mares particularly alters the
expression of genes related to transcription and RNA processing in the ICM and embryo-maternal
communication in the TE, suggesting embryo adaptation to an environment that is different in
nulliparous vs multiparous mares. Individual chances of implantation for each embryo could not be
predicted by the results of this study. Until today, the capacity of uterus to enlarge and support
pregnancy was the only suggested explanation for the lighter and smaller foal and placenta at birth in
nulliparous mares. The present results indicate differences in embryo-maternal communication long
before implantation that could alter the embryo development as well as maternal recognition of
pregnancy.
**Data Availability Statement**


**Conflict of interest**

The authors declare no conflicts of interest.

**Declaration of funding**

This work was supported by the "Institut Français du Cheval et de l'Equitation" (grant numbers CS_2018_23, 2018). The National Research Institute for Agriculture, Food and Environment (INRAE) department Animal Physiology and Breeding Systems also supported this research.

**Acknowledgments**

The authors are grateful to the staff of the Institut Français du Cheval et de l'Equitation (IFCE) experimental farm (Plateau technique de la Valade, Chamberet, France) for care and management of animals. We acknowledge the high-throughput sequencing facility of I2BC for its sequencing and bioinformatics expertise. The bioinformatics analyses were performed thanks to Core Cluster of the Institut Français de Bioinformatique (IFB) (ANR-11-INBS-0013). Many thanks to Matthias Zytnicki and Christophe Klopp for their advice on RNA-seq de novo analysis. Many thanks to Pablo Ross who kindly provided the coordinates for the XIST gene.
Author contributions

PCP obtained the funding. PCP and VD conceived the project. VD and PCP supervised the study. ED, CA, ND, NP, VD and PCP adapted the methodology for the project. ED, CA, JAR and YJ performed the experiments. CA, ND, NP and MD provided the resources. ED, LJ, YJ and RL performed data curation. ED and LJ analyzed the data. ED wrote the original draft. All authors read, revised, and approved the submitted manuscript.
List of abbreviations

DEG: differential expressed genes
DeMixT_ICM_cells: deconvoluted gene expression in ICM cells
DeMixT_TE_cells: deconvoluted gene expression in TE cells
ECM: Extracellular matrix
ERR: embryo collection rate
FDR: false discovery rate
GO BP: Gene Ontology biological process
GO: Gene Ontology
GSEA: gene set enrichment analyses
ICM: inner cell mass
ICMandTE: inner cell mass enriched hemi-embryo
ICSI: intracytoplasmic sperm injection
IL1B: Interleukin 1 beta
KEGG: Kyoto Encyclopedia of Genes and Genomes
Log2FC: log2 fold change
NES: normalized enrichment score
OM: old multiparous mares
ON: old nulliparous mares
TE: trophoblast
TE_part: pure trophoblast hemi-embryo

XIST: X inactive Specific Transcript
References


Figure legends

Fig. 1: Gene expression in ICM and TE before and after deconvolution using DeMixT

a) Venn diagram of genes differentially expressed in the different analyses: ICMandTE vs TE_part (before deconvolution), DeMixT_ICM_cells vs DeMixT_TE_cells (after deconvolution) and DeMixT_ICM_cells vs TE_part (gene expression of ICM after deconvolution vs gene expression in TE_part without deconvolution); b) Principal Component Analysis of gene expression from DeMixT_ICM_cells, DeMixT_TE_cells, ICMandTE and TE part datasets.

Deconvolution was used to isolate gene expression of ICM and TE cells in ICMandTE hemi-embryos. ICMandTE: inner cell mass + trophoblast; TE_part: pure trophoblast. Here trophoblast represents trophectoderm + endoderm.

Fig. 2: Analysis of differentially expressed genes (DEG) in embryos according to maternal parity
A) representation of down- (blue) and up- (red) regulated DEG in ICM (from DeMixT_ICM_cells data obtained after deconvolution of ICM and TE using DeMixT R package [48,49]) and TE (from TE_part dataset) of embryos from ON vs OM.

DEG: Differentially Expressed Genes (FDR < 0.05); TE: Trophoblast; ICM: Inner Cell Mass; ON: Old nulliparous mares; OM: Old multiparous mares

Fig. 3: SUMER clustering of GSEA terms clustering of the most perturbed terms in the ICM and TE of embryos according to mares’ parity

Nodes represent altered gene sets in the ICM and TE (FDR < 0.05). Node size represents the gene set size. Node shape represents the gene set database: GO BP (circle) or KEGG (diamond) or REACTOME (square). Gene sets are represented in blue if enriched (NES > 0) in young nulliparous mares’ embryos and in green if enriched (NES < 0) in young multiparous mares’ embryos. The lighter the color, the more the NES is close to 0. Edges represent the level of connection between representative gene sets. This graph was performed using SUMER R package [59] and modified using cytoscape 3.8.2 [60].

GSEA: Gene set enrichment analysis; ICM: Inner cell mass; TE: trophoblast; FDR: False Discovery Rate; GO BP: Gene Ontology Biological Process; Kyoto Encyclopedia of Genes and Genomes; NES: Normalized Enrichment Score
Supplementary material

Supplementary Figure 1:

SupFig1_Embryosize.tif

Plot of equine individual embryo according to their size

Supplementary Figure 2:

SupFig2_comp_with_without.png

Venn diagrams of differential analyses on equine embryo gene expression according to maternal parity in the inner cell mass (ICM) and the trophoblast part (TE) with or without the largest embryo (2643µm in diameter)

Supplementary Table 1:

Sup1_ICM_Diff_avecYME5.csv

Differential gene analysis using DeSeq2 in DeMixT_ICM_cells of equine embryo at Day 8 post-ovulation according to mares' parity with the large embryo

Equine ensemble ID, orthologue human Ensembl ID, Orthologue human Entrez Gene ID, gene description, normalized counts for each embryo and parameters obtained after Deseq2 analysis (log2FoldChange, pvalue and padj (after FDR correction)) of genes expressed in ICM (after gene expression deconvolution of ICMandTE using DeMixT) of YN and YM embryos

ICM: Inner cell mass; YN: young nulliparous mares; YM: young multiparous mares
Supplementary Table 2:

```
Sup2_TE_Diff_avecYME5.csv
```

Differential gene analysis using DeSeq2 in TE_part of equine embryo at Day 8 post-ovulation according to mares’ parity with the large embryo

Equine ensemble ID, orthologue human Ensembl ID, Orthologue human Entrez Gene ID, gene description, normalized counts for each embryo and parameters obtained after DeSeq2 analysis (log2FoldChange, pvalue and padj (after FDR correction)) of genes expressed in TE_part of YN and YM embryos

TE: trophoblast; YN: young nulliparous mares; YM: young multiparous mares

Supplementary Table 3:

```
Sup3_ICM_Diff_sansYME5.csv
```

Differential gene analysis using DeSeq2 in DeMixT_ICM_cells of equine embryo at Day 8 post-ovulation according to mares’ parity without the largest embryo

Equine ensemble ID, orthologue human Ensembl ID, Orthologue human Entrez Gene ID, gene description, normalized counts for each embryo and parameters obtained after DeSeq2 analysis (log2FoldChange, pvalue and padj (after FDR correction)) of genes expressed in ICM (after gene expression deconvolution of ICM and TE using DeMixT) of YN and YM embryos

ICM: Inner cell mass; YN: young nulliparous mares; YM: young multiparous mares
Supplementary Table 4:

Sup4_TE_Diff_sansYME5.csv

Differential gene analysis using DeSeq2 in TE_part of equine embryo at Day 8 post-ovulation according to mares' parity with the largest embryo

Equine ensemble ID, orthologue human Ensembl ID, Orthologue human Entrez Gene ID, gene description, normalized counts for each embryo and parameters obtained after DeSeq2 analysis (log2FoldChange, pvalue and padj (after FDR correction)) of genes expressed in TE_part of YN and YM embryos

TE: trophoblast; YN: young nulliparous mares; YM: young multiparous mares

Supplementary Table 5:

Sup5_ICM_GSEA_sansYME5.csv

Gene set enrichment analysis results on gene expression of DeMixT_ICM_cells of embryos from young nulliparous and multiparous mares

Gene Set Enrichment Analysis results (database, pathway name, size, enrichment score without and with normalization, p-value and FDR corrected q-value) for GO biological process, KEGG and REACTOME databases on DeMixT_ICM_cells gene expression table. These results did not include YM_E5, the embryo larger than 2,000 µm. ICM: Inner cell mass

Supplementary Table 6:

Sup6_TE_GSEA_sansYME5.csv

Gene set enrichment analysis results on gene expression of TE_part of embryos from young nulliparous and multiparous mares
Gene Set Enrichment Analysis results (database, pathway name, size, enrichment score without and with normalization, p-value and FDR corrected q-value) for GO biological process, KEGG and REACTOME databases on TE_part gene expression table. These results did not include YM_E5, the embryo larger than 2,000µm.
### Table 1: Mares’ characteristics at embryo collection time.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Nulliparous (YN)</th>
<th>Multiparous (YM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (n = 10)</td>
<td>With embryo (n = 6)</td>
</tr>
<tr>
<td>Breed</td>
<td>AA n = 7; SF n = 3</td>
<td>AA n = 5; SF n = 1</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>6.00 ± 0.00</td>
<td>6.00 ± 0.00</td>
</tr>
<tr>
<td>Parity (number of foalings)</td>
<td>0 ± 0.00</td>
<td>0 ± 0.00</td>
</tr>
<tr>
<td>Weight (in kg)</td>
<td>535.66 ± 30.72</td>
<td>544.55 ± 30.50</td>
</tr>
<tr>
<td>BCS (scale 1-5)</td>
<td>2.35 ± 0.32</td>
<td>2.21 ± 0.25</td>
</tr>
<tr>
<td>Withers' height (in cm)</td>
<td>159.70 ± 3.33</td>
<td>161.67 ± 2.50</td>
</tr>
</tbody>
</table>

AA: Anglo Arab or Anglo-Arabian type; SF: Selle Français section A or B; SB: Saddlebred. Age, parity, weight, and height are presented as mean ± SD.
Table 2: Comparison of the expression of selected genes previously identified as specific to TE or ICM in equine embryos [51], before and after deconvolution

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Ensembl ID</th>
<th>ICMandTE vs TE_part</th>
<th>DeMixT_ICM_cells vs TE_part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>log2FC from DeSeq2</td>
<td>padj</td>
</tr>
<tr>
<td>SOX2</td>
<td>ENSECA600000010653</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NANO2</td>
<td>ENSECA600000012614</td>
<td>5.78</td>
<td>6.93E-58</td>
</tr>
<tr>
<td>SPP1</td>
<td>ENSECA600000017191</td>
<td>4.52</td>
<td>2.21E-12</td>
</tr>
<tr>
<td>LIN28B</td>
<td>ENSECA600000020994</td>
<td>3.21</td>
<td>6.84E-13</td>
</tr>
<tr>
<td>SMARCA2</td>
<td>ENSECA600000024187</td>
<td>1.01</td>
<td>0.139</td>
</tr>
<tr>
<td>POU5F1 (OCT4)</td>
<td>ENSECA600000008967</td>
<td>0.62</td>
<td>8.32E-04</td>
</tr>
<tr>
<td>ID2</td>
<td>ENSECA600000008738</td>
<td>0.33</td>
<td>0.604</td>
</tr>
<tr>
<td>DNMT3B</td>
<td>ENSECA600000012102</td>
<td>0.49</td>
<td>0.056</td>
</tr>
<tr>
<td>DPPA4</td>
<td>ENSECA600000013271</td>
<td>0.39</td>
<td>0.545</td>
</tr>
<tr>
<td>SALL4</td>
<td>ENSECA600000018533</td>
<td>0.21</td>
<td>0.177</td>
</tr>
<tr>
<td>KLF4</td>
<td>ENSECA600000010613</td>
<td>0.03</td>
<td>0.995</td>
</tr>
<tr>
<td>UTF1</td>
<td>ENSECA600000039888</td>
<td>0.34</td>
<td>0.726</td>
</tr>
<tr>
<td>TFAP2A</td>
<td>ENSECA600000017468</td>
<td>-0.21</td>
<td>0.010</td>
</tr>
<tr>
<td>CDX2</td>
<td>ENSECA600000027754</td>
<td>-0.21</td>
<td>0.290</td>
</tr>
<tr>
<td>ELF3</td>
<td>ENSECA600000014608</td>
<td>-0.12</td>
<td>0.864</td>
</tr>
<tr>
<td>GATA2</td>
<td>ENSECA600000016768</td>
<td>-0.10</td>
<td>0.913</td>
</tr>
<tr>
<td>GATA3</td>
<td>ENSECA600000024574</td>
<td>-0.18</td>
<td>0.28</td>
</tr>
<tr>
<td>TEAD4</td>
<td>ENSECA600000011303</td>
<td>-0.19</td>
<td>0.206</td>
</tr>
<tr>
<td>FREM2</td>
<td>ENSECA600000020410</td>
<td>-0.06</td>
<td>0.970</td>
</tr>
</tbody>
</table>

Gene expressions were obtained from RNA of 11 equine embryos bissected in two hemi-embryos: one part was composed only of trophoblast (TE), TE_part, while the other part was composed of TE and inner cell mass (ICM), ICMandTE. As it is impossible to estimate the proportion of each cell in ICMandTE, a deconvolution algorithm (package DeMixT) was used to estimate gene expression of these different kind of cells. DeMixT_ICM_cells dataset corresponds to the deconvoluted gene
expression of ICM cells from ICM and TE. Log2 fold change (log2FC) and padj (adjusted p-value with Benjamini-Hochberg correction) were obtained with Deseq2 package. TE_part is the reference group in both analyses: when log2 fold changes (log2FC)>0, the gene is more expressed in the ICM and TE or DeMixT_ICM_cells, while when log2FC<0, the gene is more expressed in the TE_part.

Green is used to represent genes differentially expressed in the present study. Orange is used to represent genes that have been previously identified as predominant in the ICM [51] but that are identified here as predominant in the TE.
Table 3: Up- and down-regulated genes coding for a protein in the inner cell mass of equine embryos according to mare parity

<table>
<thead>
<tr>
<th>Ensembl Name</th>
<th>Entrez Gene ID</th>
<th>Description</th>
<th>GO Molecular function</th>
<th>GO Biological Process</th>
<th>log2 Fold Change</th>
<th>padj</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSECAG00000029895</td>
<td>MAGEB16</td>
<td>MAGE family member B16</td>
<td>Tumor antigen</td>
<td>Extracellular matrix disassembly Positive regulation of cell killing Regulation of hydrogen peroxide metabolic process Superoxide metabolic process Regulation of respiratory burst</td>
<td>-3.24</td>
<td>0.013</td>
</tr>
<tr>
<td>ENSECAG00000017619</td>
<td>NOX1</td>
<td>NADPH oxidase organizer 1</td>
<td>Phospholipid binding Superoxide-generating NADPH oxidase activator activity</td>
<td>Extracellular matrix disassembly Positive regulation of cell killing Regulation of hydrogen peroxide metabolic process Superoxide metabolic process Regulation of respiratory burst</td>
<td>-2.67</td>
<td>0.034</td>
</tr>
<tr>
<td>ENSECAG00000023392</td>
<td>ZBTB8A</td>
<td>zinc finger and BTB domain containing 8A</td>
<td>Metal ion binding DNA binding</td>
<td>Regulation of transcription by RNA polymerase II Transcription regulation</td>
<td>-2.02</td>
<td>5.84E-05</td>
</tr>
<tr>
<td>ENSECAG00000019702</td>
<td>VPS52</td>
<td>VPS52 subunit of GARP complex</td>
<td>Syntaxin binding</td>
<td>Ectodermal cell differentiation Embryonic ectodermal digestive tract development Endocytic recycling Lysosomal transport Protein transport</td>
<td>-1.26</td>
<td>0.001</td>
</tr>
<tr>
<td>ENSECAG00000011960</td>
<td>DES1</td>
<td>desumoylating isopeptidase 1</td>
<td>Hydrolase Identical protein binding</td>
<td>Protein desumoylation Protein export from nucleus Protein modification by small protein removal</td>
<td>-1.23</td>
<td>0.001</td>
</tr>
<tr>
<td>ENSECAG00000020433</td>
<td>TENM3</td>
<td>teneurin transmembrane protein 3</td>
<td>Cell adhesion molecule binding Protein heterodimerization activity Protein homodimerization activity</td>
<td>Cell adhesion Differentiation Neuron development Signal transduction</td>
<td>-1.00</td>
<td>0.015</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Name</td>
<td>Function</td>
<td>ENSECA00000015867</td>
<td>ENSECA0000004931</td>
<td>ENSECA00000034815</td>
<td>ENSECA0000007262</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>PAQR3</td>
<td>progestin and adipQ receptor family member 3</td>
<td>Signaling receptor activity</td>
<td>-0.84</td>
<td>0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARS2</td>
<td>methionyl-tRNA synthetase 2, mitochondrial</td>
<td>Aminoacyl-tRNA synthetase Ligase ATP binding</td>
<td></td>
<td>-0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VHL</td>
<td>Von Hippel-Lindau tumor suppressor</td>
<td>Enzyme binding Transcription factor binding</td>
<td></td>
<td></td>
<td>-0.81</td>
<td></td>
</tr>
<tr>
<td>EEA1</td>
<td>early endosome antigen 1</td>
<td>GTP-dependant protein binding Protein homodimerization activity</td>
<td></td>
<td></td>
<td></td>
<td>-0.79</td>
</tr>
<tr>
<td>GSTCD</td>
<td>glutathione S-transferase C-terminal domain containing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Gene</td>
<td>Description</td>
<td>Functions</td>
<td>Fold-Change</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>-------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000014243</td>
<td>GATM</td>
<td>Glycine amidinotransferase activity</td>
<td>Creatine metabolic process</td>
<td>0.73</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>ENSECAG000000041817</td>
<td>RBP1</td>
<td>Retinol binding protein 1 activity</td>
<td>Lipid homeostasis</td>
<td>1.13</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000010447</td>
<td>EFEMP1</td>
<td>EGF containing fibulin extracellular matrix protein 1</td>
<td>Embryonic eye morphogenesis</td>
<td>1.95</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000010385</td>
<td>MET</td>
<td>MET proto-oncogene, receptor tyrosine kinase</td>
<td>Branching morphogenesis of an epithelial tube</td>
<td>2.83</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000022277</td>
<td>MAOA</td>
<td>Monoamine oxidase A</td>
<td>Cellular biogenic amine metabolic process</td>
<td>4.51</td>
<td>0.001</td>
<td></td>
</tr>
</tbody>
</table>

Log2 Fold-Change<0 indicates down-regulation of the gene in embryos from nulliparous mares, also indicated in green; Log2 Fold-Change>0 indicates up-regulation of the gene in embryos from nulliparous mares, also indicated in blue.
Table 4: Up- and down-regulated genes coding for a protein in the trophoblast part of equine embryos according to mare parity

<table>
<thead>
<tr>
<th>Ensembl Name</th>
<th>Ensemble Name</th>
<th>Description</th>
<th>GO Molecular function</th>
<th>GO Biological Process</th>
<th>log2 Fold Change</th>
<th>padj</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSECAG00000008177</td>
<td>PRXL2B</td>
<td>peroxiredoxin like 2B</td>
<td>Antioxidant activity</td>
<td>Prostaglandin biosynthetic process</td>
<td>-2.15</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Prostaglandin-F synthase activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSECAG000000012493</td>
<td>SLC47A1</td>
<td>solute carrier family 47 member 1</td>
<td>Amide transmembrane transporter activity</td>
<td>Amino acid import across plasma membrane</td>
<td>-1.47</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L-amino acid transmembrane transporter activity</td>
<td>L-arginine import across plasma membrane</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Organic cation transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000018727</td>
<td>EFCA11</td>
<td>EF-hand calcium binding domain 11</td>
<td>Calcium ion binding</td>
<td>Cellular response to hypoxia</td>
<td>0.51</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive regulation of apoptotic process</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive regulation of release of cytochrome c from mitochondria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000020415</td>
<td>FAM162A</td>
<td>family with sequence similarity 162 member A</td>
<td>GTP binding</td>
<td>GTP binding</td>
<td>0.65</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Metal ion binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENSECAG00000012554</td>
<td>GTPBP8</td>
<td>GTP binding protein 8 (putative)</td>
<td></td>
<td></td>
<td>0.93</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log2 Fold-Change<0 indicates down-regulation of the gene in embryos from nulliparous mares, also indicated in green; Log2 Fold-Change>0 indicates up-regulation of the gene in embryos from nulliparous mares, also indicated in blue.