Biology of Reproduction, 2021, 105(6), 1416–1426
https://doi.org/10.1093/biolre/ioab174
Research Article
Advance Access Publication Date: 13 September 2021

Research Article

Allele-specific expression analysis reveals conserved and unique features of preimplantation development in equine ICSI embryos†

D.E. Goszczynski[1], P.S. Tinetti[2], Y.H. Choi[2], P.J. Ross[1,‡] and K. Hinrichs[2,‡,*]

[1]Department of Animal Science, University of California, Davis, CA 95616, USA. and [2]Department of Veterinary Physiology and Pharmacology, Texas A&M University, College Station, TX 77843, USA.

***Correspondence**: Department of Clinical Studies – New Bolton Center, 382 W Street Rd, Kennett Square, PA 19348. E-mail: katrinh@vet.upenn.edu (Katrin Hinrichs); 2239 Meyer Hall, Davis, CA 95616, USA. E-mail: pross@ucdavis.edu (Pablo Juan Ross)

†Grant Support: Research in Dr. Hinrichs' laboratory was supported by the Clinical Equine ICSI program at Texas A&M University. D.E. Goszczynski and P.S. Tinetti were Fulbright Scholars during their stay at the Ross and Hinrichs laboratories.
‡Joint senior authors

Received 30 March 2021; Revised 17 August 2021; Accepted 10 September 2021

Abstract

Embryonic genome activation and dosage compensation are major genetic events in early development. Combined analysis of single embryo RNA-seq data and parental genome sequencing was used to evaluate parental contributions to early development and investigate X-chromosome dynamics. In addition, we evaluated dimorphism in gene expression between male and female embryos. Evaluation of parent-specific gene expression revealed a minor increase in paternal expression at the 4-cell stage that increased at the 8-cell stage. We also detected eight genes with allelic expression bias that may have an important role in early development, notably NANOGNB. The main actor in X-chromosome inactivation, XIST, was significantly upregulated at the 8-cell, morula, and blastocyst stages in female embryos, with high expression at the latter. Sexual dimorphism in gene expression was identified at all stages, with strong representation of the X-chromosome in females from the 16-cell to the blastocyst stage. Female embryos showed biparental X-chromosome expression at all stages after the 4-cell stage, demonstrating the absence of imprinted X-inactivation at the embryo level. The analysis of gene dosage showed incomplete dosage compensation ($0.5 < X{:}A < 1$) in MII oocytes and embryos up to the 4-cell stage, an increase of the X:A ratio at the 16-cell and morula stages after genome activation, and a decrease of the X:A ratio at the blastocyst stage, which might be associated with the beginning of X-chromosome inactivation. This study represents the first critical analysis of parent- and sex-specific gene expression in early equine embryos produced in vitro.

Summary sentence

Early development of equine ICSI-derived embryos involves activation of the paternal genome at the 4-cell stage; genes with allele-specific expression bias; sexual dimorphism in gene expression with strong representation of the X-chromosome; expression of XIST in both sexes since genome

activation with increasingly higher expression in late female embryos; absence of parent-specific X-inactivation at the embryo level; and scarce dosage compensation until the blastocyst stage.

Key words: horse, genome activation, dosage compensation, X-chromosome inactivation, allelic bias, ICSI, sexual dimorphism, gene expression.

Introduction

Embryonic genome activation (EGA) is the key event in early embryo development, during which transcription from the new embryonic genome begins. This process is dictated by maternal factors stored in the oocyte during oogenesis, such as mRNA transcripts, which are translated by the embryo after fertilization [1]. However, the timing and mechanisms of EGA differ between species, with extensive de novo transcription occurring at the 2-cell stage in mice [2], at the 4-to-8-cell stage in humans [3], cattle [4], and pigs [5], and at the 8-cell stage in macaques [6]. Even the origin of the embryo can be determinant, as cloned pig embryos activate their genome one stage later than in vivo embryos [5].

In horses, very little is known about EGA, with early studies reporting signs of transcription at the 5-cell [7] and 6-cell [8] stages. The horse is an economically important species that represents a separate family of mammals (Perissodactyla), together with rhinos and tapirs, which are endangered species. Therefore, the study of key events in the early development of equine embryos would not only contribute to the improvement of assisted reproductive technologies for breeding or veterinarian purposes, but also serve as a base for comparative studies for the conservation of endangered species. Moreover, the horse represents an interesting model for the study of the placenta due to their epitheliochorial and minimally invasive placentation, which unlike murine and human placentation, allows for distinction of embryonic transcripts from maternal components at the fetomaternal interface. The study of gene expression and parental contribution to preimplantation development would fill the knowledge gap between fecundation and implantation.

As embryonic transcripts can be confounded with the highly abundant maternal mRNAs provided by the oocyte during the early developmental stages, the initiation of embryonic transcription can be identified through the increase of transcripts carrying paternal alleles. Parental genome information is also useful to evaluate parental contribution to embryonic development. In mammals, both the paternal and maternal genomes are required for normal embryogenesis due to differences in genomic imprinting [9, 10]. Differential methylation in the parental genomes frequently derives in allele-specific expression (ASE), which is well studied in adult tissues of model species but poorly understood in preimplantation embryos. Approximately 1400 genes with parent-of-origin expression bias have recently been reported in human embryos [11] and horse placenta [12]. However, the expression degree of imprinted genes varies among species and tissues; for example, one study found no evidence of active demethylation of the paternal genome in ICSI equine embryo pronuclei [13], similar to sheep and rabbits, but unlike mice, pigs, and humans. In the horse, ~1300 genes have been reported as parentally biased in the placenta [12], but one study has shown that many of the parentally-biased genes in the placenta presented no bias in the fetus [14]. To the best of our knowledge, there are no reports of parentally biased genes in equine preimplantation embryos.

Another key event in early development is dosage compensation. In mammals, maintenance of proper gene dosage is essential for embryo survival. Aneuploidies frequently lead to developmental disorders and abortions [15], and even changes in the expression of single genes can cause tumorigenesis and disease [16]. Nevertheless, males tolerate the absence of a second X-chromosome. The need to balance gene dosage between male and female is addressed by silencing of the majority of one of the X-chromosomes in each female cell by the complex process of X-chromosome inactivation (XCI); the remainder of the silenced X-chromosome escapes XCI and continues to be active as a pseudo-autosomal region, homologous to the pseudo-autosomal region of the Y chromosome [17]. XCI has been reported in several species, including the horse [18]. Mechanisms of XCI, and the developmental stage at which it occurs, differ between species. In mouse embryos, there are two types of XCI that ensure the silencing of one of the X chromosomes in a stage-specific manner: At the 4-cell stage, a skewed inactivation of the paternal X-chromosome occurs, associated with a maternal *XIST*-repressing imprint [19]; this is followed by a brief stage in which both chromosomes are active, and then later in development, the inner cell mass switches to random XCI [20]. Paternal X-chromosome silencing has also been reported in bovine embryos at the morula stage [21] and in marsupial embryos [22, 23]. In other species such as human and rabbit, XCI appears to be random through embryo development [20]. In equids, XCI was studied via RNA-seq in chorionic girdle from horse and donkey-horse hybrid fetuses, and the authors concluded that XCI is random [24]. An early study utilizing 5-bromodeoxyuridine incorporation in equine conceptuses recovered by uterine flush, found that XCI began gradually in the extraembryonic cells around Day 7.5 and in the embryonic disc around Day 11.5 [18]. However, the processes of XCI in early equine embryos have never been evaluated from a transcriptomic perspective.

The objective of this study was to evaluate the parental contribution to preimplantation equine embryo development via analysis of allele-specific expression. For this purpose, we complemented RNA-seq data from equine ICSI-derived embryos with whole-genome DNA-Seq data from the stallion and mares used to produce those embryos. The stallion used for these studies was a Norwegian Fjord Horse, a breed identified as being genetically dissimilar to the Quarter-type mares utilized [25] and thus likely to provide a high number of single-nucleotide polymorphisms to allow parental-specific gene identification, while avoiding the use of interspecific hybrid embryos (e.g. horse-donkey hybrids). This complementation allowed us to track the initiation of EGA by paternal expression, to evaluate allele-specific expression, and to investigate X-chromosome dynamics during preimplantation development.

Materials and methods

Animal samples

Three Quarter Horse-type mares (10–15 years old) were group-housed in pens with *ad libitum* access to hay and fresh water. Oocytes were collected from these mares by transvaginal ultrasound-guided aspiration and blood samples were collected for whole genome sequencing. All procedures were approved by the Texas A&M

University Institutional Animal Care and Use Committee and were carried out in accordance with the *United States Government Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research and Training* (Protocol 2015–0282). Semen from one ejaculate of a 14-year-old Norwegian Fjord Horse stallion was collected, extended, and shipped cooled to the laboratory. Once at the laboratory, the semen was frozen using a standard freezing protocol for equine semen with E-Z Freezin-MFR5 medium (Animal Reproduction Systems, Chino, CA).

Experimental design

Oocytes from the three mares (D, J, and S) were matured *in vitro*, denuded, and fertilized with sperm using intracytoplasmic sperm injection (ICSI). The study was conducted using oocytes/embryos in each of eight stages: unfertilized metaphase-II (MII) oocytes, zygote, 2-cell, 4-cell, 8-cell, 16-cell, morula, and blastocyst. The dataset consisted of six embryos per stage (two per mare), except for the morula stage, which contained three replicates from Mare D (D1, D2, D3) and one replicate from Mare S (S2) due to a case of mislabeling identified after the analysis of paternal expression.

Chromatin/nuclear status was confirmed for each oocyte and embryo used in the study by evaluation of the embryo after staining with Hoechst 33342. Embryos were selected if the number of nuclei corresponded to the observed number of cells and no morphological evidence of cellular or chromatin degeneration were found. Selected mature oocytes presented normal MII chromatin configuration (metaphase plate and polar body). Embryos classified as zygotes had two pronuclei and two polar bodies; morulae had >32 nuclei. Blastocysts had >64 nuclei and an organized outer ring of nuclei. Selected oocytes/embryos were placed into 0.5 ml RNase-free microfuge tubes (Ambion, LifeTechnologies) with < 0.5 µl of medium, snap-frozen in liquid nitrogen, and stored at −80°C until RNA extraction.

DNA-seq data generation and processing

Mare blood samples and stallion semen were sent to the UC Davis DNA Technologies Core for whole-genome sequencing. DNA was sequenced as paired reads (150 bp) in an Illumina HiSeq4000 platform. All samples were sequenced at a minimum average genome coverage of 17X. Reads were processed following the best practices for data pre-processing and variant discovery in GATK v4.1.4.1 [26]. Briefly, unmapped BAM files containing the metadata were generated using fastqToSam, adapters were marked using the MarkIlluminaAdapters tool, reads were aligned to the EquCab3.0 genome using the MEM procedure of BWA v0.7.13 [27], and lastly, unmapped and mapped BAM files were merged using MergeBamAlignment. Duplicates were marked using Picard Tools v2.8.1 (broadinstitute.github.io/picard). Individual GVCF files were generated using HaplotypeCaller and combined using CombineGVCFs. Then, variants were called from the combined GVCF using GenotypeGVCFs. Variants were filtered by parameters determined empirically (Supplemental Figure S1) to keep only those with high quality for recalibration: QD < 5, FS > 12, SOR < 3, MQ < 59, MQRankSum <−0.3, ReadPosRankSum<−8. Base quality scores were recalibrated in the BAM files using the BaseRecalibrator and ApplyBQSR tools and considering variants from the previous step as "known" variants. Variants were called a second time in the recalibrated dataset using HaplotypeCaller, CombineGVCFs, and GenotypeGVCFs, and filtered using the same criteria as in the first round, except for a more stringent QD cutoff (10) (Supplemental Figure S2). Lastly, indels were discarded, and only biallelic SNPs with a minimum depth of 10 reads were retained in the dataset. From the millions of parental SNPs, only those with informative genotypes for the allele-specific expression analysis of each mare-stallion pair were retained in the dataset (mare AA – stallion BB; mare AA – stallion AB). In other words, only the sites that allowed the identification of the paternal allele were retained.

RNA-Seq library preparation and data processing

Stranded RNA-Seq libraries were constructed from each embryo using the Ovation Solo RNA-Seq System (NuGen Techonology, San Carlos, CA) following the manufacturer's protocol. The number of cycles required for library amplification was determined for each sample by the exponential phase of amplification using EvaGreen 20X (Biotium) in a qPCR. The number of amplification cycles varied from 11 to 20 cycles (94°C 30s, 60°C 30s, 72°C 1 m). Amplified cDNA was purified using AMPure XP Beads (Beckman Coulter) and the concentration was evaluated using a Qubit DNA HS assay (Invitrogen, Carlsbad, CA). Then, a second round of amplification (2 cycles, 95°C 30s, 60°C 90s; 6 cycles, 95°C 30s, 65°C 90s) and purification with beads was carried out. Quantity and quality of final libraries were assessed using a High Sensitivity DNA Bioanalyzer chip (Agilent Technologies, Santa Clara, CA) using approximately 2–5 ng DNA per sample (Supplemental File S1). Libraries sized between 200 and 1000 bp and were sequenced as paired end reads of 100–150 bases in an Illumina Hi-Seq 4000 apparatus located at the UC Davis DNA Technologies Core.

Raw sequences were aligned to a reference genome composed of the recently assembled Y-chromosome [28] and the v97 ENSEMBL release of the EquCab3 genome. Alignments were carried out using STAR v2.7.0f [29] in two-pass mode with WASP filtering. In the WASP procedure, alleles from reads that overlap with SNPs are replaced by the alternative alleles, and reads are remapped to the genome. If a remapped read does not exactly match the same locus, the aligner flags the read for future removal due to potential mapping bias. Duplicates were marked using Picard Tools and only uniquely mapped reads were retained. Gene expression was quantified using the featureCounts tool from the Subread package v1.6.2 [30] in stranded mode, ignoring duplicates, assigning reads to features based on the largest overlap, and counting fragments with both ends aligned to the same strand of the same chromosome. Official gene symbols, and human homologs (when no external gene names were available), were collected using the biomaRt v2.40.5 R package [31]. Embryo sexing was carried out by expression of ancestral Y-chromosome genes [28]. Differential gene expression between male and female embryos was evaluated using the DESeq2 R package [32]. Evaluations were carried out using Wald tests and the v97 ENSEMBL release of the EquCab3 annotation, with manual inclusion of the *XIST* gene. As male and female embryos have been reported to present small fold changes in differentially expressed genes in several species, we used an adjusted p-value of 0.1 as significance cut-off. Gene ontology analyses were carried out using the Functional Annotation tool from DAVID Bioinformatics [33] with the "GOTERM_MF_ALL", "GOTERM_BP_ALL", and "GOTERM_CC_ALL" categories. Since the human GO database is more refined than horse, we used human background in the analysis. The most relevant GO terms were plotted using the GOplot R package [34].

Quantification of allele-specific expression

Allele counts were determined for each embryo at the informative sites identified by comparing dam and sire (D-stallion, J-stallion,

S-stallion) sequences, using the allele-specific expression ASEReadCounter tool from the GATK package. The ASEReadCounter tool quantifies ASE by site, based on the number of reads carrying each allele in a list of sites provided as a VCF file. As this tool only counts heterozygous sites by default (i.e. mare AA – stallion BB), custom VCF files were generated to guide the tool into counting sites for which, based on sire and dam genotypes, both homozygous and heterozygous combinations were possible (i.e. mare AA – stallion AB/mare AA – stallion BB), for embryos that received a B paternal contribution.

Identification of the timing of genome activation

To determine the timing of EGA by the onset of paternal expression, informative heterozygous exon sites were filtered by expression in each sample (total counts ≥10), and paternal ratios (PR) for each gene were calculated as paternal expression over total expression. When genes presented multiple SNPs, the counts from the individual sites were summed to calculate the PR. Only autosomal genes expressed in at least two biological replicates (total counts ≥10 in each sample) of a given stage were considered for this step.

Evaluation of allelic expression bias

Parent-of-origin allelic expression bias was evaluated at stages posterior to EGA. Genes with potential allelic bias were identified by evaluating the PR at each individual site. If a gene presented multiple informative sites with sufficient counts (≥ 10), the consistency of this bias was evaluated across the remaining sites. Genes were considered biased when the trend was observed in all the biological replicates of a given stage that counted with informative genetic variation and sufficient counts (≥ 10). Once identified, genes were assigned a PR value calculated as the sum of paternal counts over the total number of counts at all informative sites. Based on their PR, genes with allelic bias were classified as paternally biased (PR ≥ 0.75) or maternally-biased (PR ≤ 0.25).

Evaluation of X-chromosome dynamics

The existence of imprinted X-inactivation was evaluated based on the presence of reads carrying the paternal and maternal alleles at each informative site identified on the X-chromosome. Sites presenting heterozygous genotypes in the X-chromosome of the stallion were filtered out to remove mapping errors. To calculate X:A ratios and evaluate dosage compensation, we used the pairwiseCI R package to obtain 95% confidence intervals for the ratios of the median of X to the median of A. The confidence intervals were calculated using bootstrapping (10,000 replicates), which involved random sampling from a distribution with replacement. In each sample, only genes with at least 4 transcripts per million (TPM) were considered to calculate the X:A ratio. All downstream analyses were carried out using R, BCFtools v1.10.2 [35] and VCFtools v0.1.14 [36].

Results

Genetic variation in parents and embryos

Analysis of whole genome sequencing identified approximately 4 million biallelic SNP sites in the stallion and each of the mares, with average genome coverages of at least 17X (Table 1). About 2.5 million of these sites were informative for the purposes of allele-specific expression, i.e. the mare presented an AA genotype while the stallion presented a BB (or AB) genotype, allowing the assignment of the B allele to the stallion. As expected, most sites were in intergenic regions, and ~ 38,000 (1%) of them were in annotated exons; these corresponded to approximately 14,000 genes.

RNA-Seq libraries yielded 1.4 billion uniquely mapped reads, with an average of 29 million reads per sample and an average unique mapping rate of 80%. Of these 29 million uniquely mapped reads per embryo, ~845,000 were overlapped informative sites that passed the WASP filter (unbiased mapping). A hierarchical clustering analysis based on variance stabilized counts showed distinct clusters corresponding to each developmental stage after the 2-cell stage, except for the "4cell-D2" sample, which showed a delayed profile and was discarded form the study. Oocytes, zygotes, and 2-cell embryos clustered together. The main limitation of our dataset relied in the high abundance of duplicate reads, associated with the limited amount of starting RNA, which drastically reduced the number of useful reads overlapping informative sites to ~134,000 per embryo. RNA-Seq library complexity allowed us to detect ~12,000 informative sites (AA-BB, AA-AB) in exons with different expression levels (counts ≥1; Supplemental File S2), corresponding to ~5000 genes per mare. Considering ten expression counts as minimum to determine parental contribution at a given site, approximately 2600 informative exon sites (AA-BB, AA-AB), corresponding to ~1400 genes on average, were detected in at least one embryo transcriptome from each mare (mare S: 1.96 SNPs/gene; mare J: 1.84 SNPs/gene; mare D: 1.88 SNPs/gene). However, many of these genes were sufficiently expressed (counts ≥10) in few replicates. The average number of strictly heterozygous sites with sufficient expression counts was of 238 sites per embryo (~161 genes). Even though the final number of informative reads per embryo was low to evaluate all genes with biased allelic expression, especially with the high variation among samples, it was sufficient to evaluate global genomic events, such as EGA and X-chromosome expression, aspects that were independent of allele-specific expression such as differential gene expression between male and female, and allelic bias in a few genes with sufficient expression counts (≥ 10 in each sample).

Evaluation of genome activation by assessment of paternal allele-specific expression

Allele information from parental genomes allowed us to determine the timing of embryonic genome activation by onset of expression of paternal-specific transcripts. To validate the methodology, we evaluated the percentage of mitochondrial reads assigned to the maternal allele in informative loci with more than 10 counts in each sample. As a result, we detected high maternal assignment (>96%) in all the embryos (Supplemental Figure S3).

We observed the first massive increase of paternal expression at the 4-cell stage, followed by a larger increase at the 8-cell stage (Figure 1A). These increases in paternal expression were consistent with the two-wave pattern described in other species and coincided with minor and major genome activation observed by evaluation of differential expression between stages (data not shown). In fact, 80% (16/20) and 91% (73/80) of the genes significantly upregulated at the 4 and 8-cell stages that presented more than ten counts showed paternal allele expression (PR > 0.1). On the other hand, only one of the eleven downregulated genes at the 8-cell stage was expressed from the paternal copy, supporting the hypothesis that this downregulation might represent maternal transcript degradation.

Once the activated stages were identified, we evaluated parent-of-origin allelic expression bias at these stages. Even though the

Table 1. Genetic variation detected in the parental genomes and embryo transcriptomes. The number of genes to which each SNP group corresponds is indicated in parentheses

	Mare S	Mare J	Mare D	Stallion
Biallelic SNPs (depth > 10)	4,112,284	4,054,016	4,590,767	4,644,961
In chromosomes (1–31, X)	4,085,216	4,029,368	4,560,914	4,617,627
Informative (AA-BB, AA-AB)	2,664,273	2,415,540	2,444,285	
Informative (AA-BB, AA-AB) in exons	38,589 (14,541)	36,805 (13,728)	38,095 (13,923)	
Informative (AA-BB, AA-AB) in at least one embryo (depth ≥ 1)	8825 (3744)	12,564 (5598)	12,580 (5453)	
Informative (AA-BB, AA-AB) in at least one embryo (depth ≥ 10)	1902 (975)	3203 (1746)	2576 (1372)	
Informative (AA-BB, AA-AB) per embryo (depth ≥ 10)	467 (261)	986 (596)	664 (391)	
Informative (AA-BB)	708,163	610,104	585,356	
Informative (AA-BB) in exons	10,382 (4846)	9069 (4365)	8916 (4176)	
Informative (AA-BB) in exons in at least one embryo (depth ≥ 1)	3383 (1777)	3946 (2107)	3629 (1868)	
Informative (AA-BB) in exons in at least one embryo (depth ≥ 10)	753 (479)	1066 (661)	745 (455)	
Informative (AA-BB) per embryo (depth ≥ 10)	186 (128)	331 (224)	193 (129)	

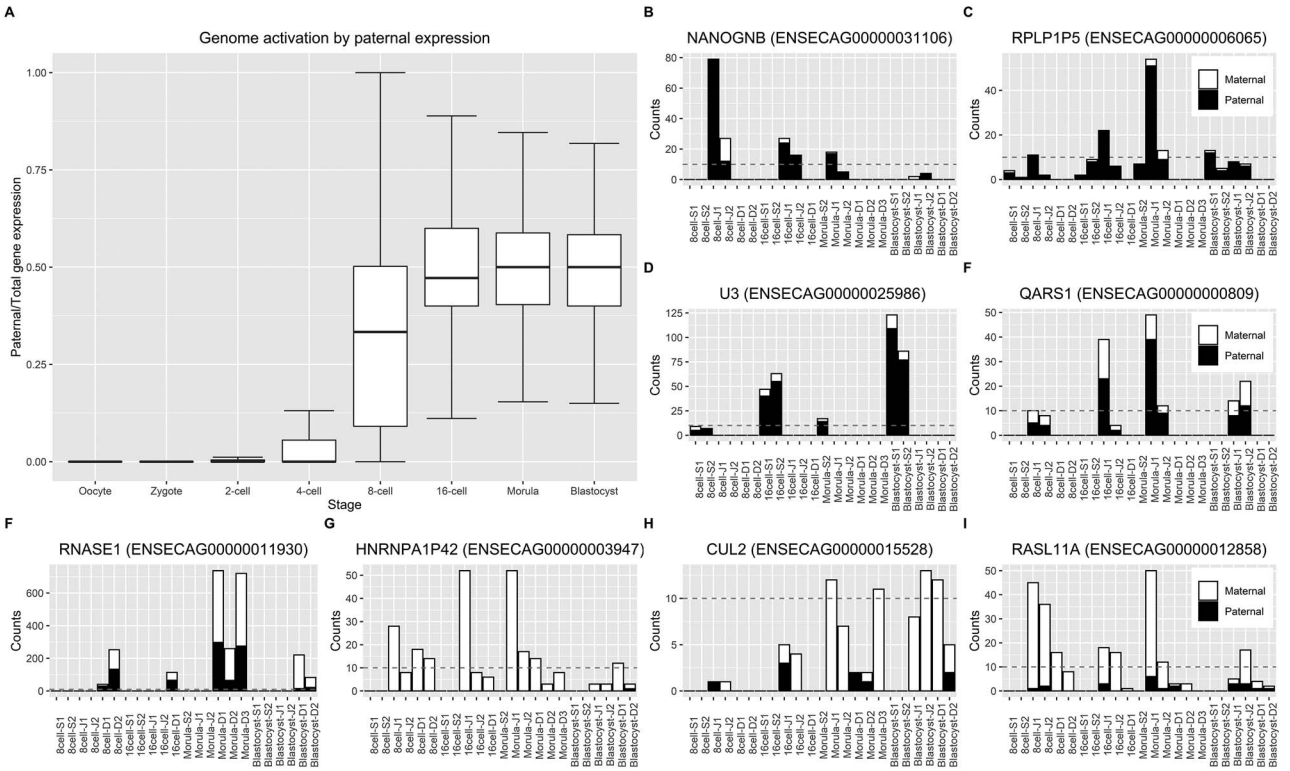

Figure 1. Allele-specific expression in horse ICSI embryos. (A) Identification of the timing of genome activation by onset of paternal gene expression. A minor increase in paternal expression was detected at the 4-cell stage, with a massive increase at the 8-cell stage. (B, C, D, E) Genes exhibiting paternal bias in all replicates of a given stage counting with ≥10 reads and informative genetic variation. (F, G, H, I) Genes similarly exhibiting maternal bias. Counts from individual sites located in the same gene were summed for purposes of visualization. In the case of genes with multiple informative loci (NANOGNB, U3, RNASE1), the loci presented consistent expression bias across the gene body, with the exception of the RNASE1 gene, which presented a split pattern (Supplemental Figure S4).

number of loci with ten or more expression counts in each sample was low, it was sufficient to evaluate allelic expression bias in a few genes with sufficient expression counts. However, the evaluation of multiple genes with allelic bias reported in human embryos [11], mouse embryonic tissues [37], and horse placenta [12, 14] resulted in either low expression counts, lack of informative genetic variation, or biparental expression in our dataset.

On evaluation of the 16-cell to blastocyst-stage embryos, we found four genes with paternal bias (*NANOGNB*, *RPLP1P5*, *U3*, *QARS1*), and four genes with maternal bias (*RNASE1*, *HNRNPA1P42*, *CUL2*, *RASL11A*) at different stages: *NANOGNB* presented paternal bias at the 16-cell stage (Figure 1B); *RPLP1P5* presented paternal bias at all stages (Figure 1C); ENSECAG00000025986, which encodes for one of the U3 snoRNAs, presented paternal bias at the 16-cell and blastocyst stages (Figure 1D); and *QARS1* presented paternal bias at the morula stage (Figure 1E). Regarding maternal bias, *RNASE1* presented bias at the blastocyst stage (Figure 1F); *HNRNPA1P42* presented bias at all stages

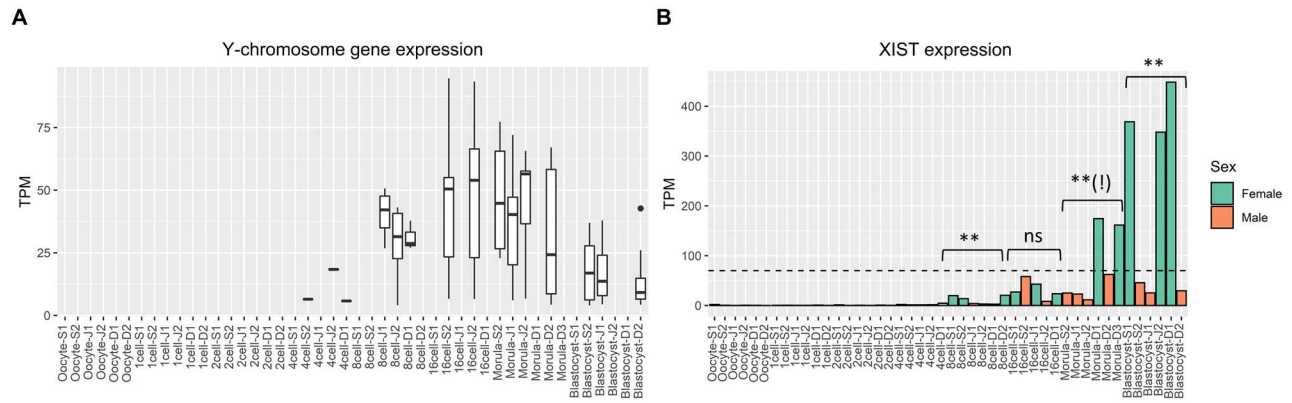

Figure 2. Y-chromosome and *XIST* expression in preimplantation horse embryos produced by ICSI. (A) Y-chromosome genes were expressed as early as the 4-cell stage, allowing male embryo identification from the 8-cell stage onward. (B) *XIST* expression in embryos designated male (orange) and female (green) based on Y-chromosome gene expression. Female embryos showed significantly higher expression than did male embryos at the 8-cell and blastocyst stages (Padj <0.01). It is worth mentioning that the Wald test corresponding to the morula stage indicated that the difference was significant ($P < 0.01$), although it could have been a false positive (Padj = 0.13).

(Figure 1G); *CUL2* presented bias at the morula and blastocyst stages (Figure 1H); and *RASL11A* presented bias at the 8-cell, 16-cell and morula stages (Figure 1I). The *NANOGNB* and *U3* genes presented two and four informative loci, respectively, and they all followed the same trend of paternal bias (Supplemental Figure S4). The *RNASE1* gene, on the other hand, presented multiple informative SNPs (13) that comprised two main regions (Supplemental Figure S5). The first region (SNPs 1–5) presented no allelic bias, while the second region (SNPs 7–13) presented maternal bias (Supplemental Figure S4).

Interestingly, we found no paternal allele expression in the genes of the 16-cell-D2 sample (Supplemental Figure S6). Based on this observation, we hypothesized the sample could be a case of parthenogenetic activation. Due to this anomalous behavior, the sample was discarded from further evaluations.

Embryo sexing by Y-chromosome and XIST gene expression

To evaluate X-chromosome inactivation and dosage compensation, we first quantified gene expression (Supplemental File S3) to determine the sex of the embryos. The first approach was to assess the expression of Y-chromosome genes. In this analysis, we found that five of the twenty-five annotated single-copy ancestral Y-genes (*UTY*, *USP9Y*, *OFD1Y*, *TMSB4Y*, *E1F1AY*) were expressed in unfertilized oocytes, suggesting they were either part of pseudo-autosomal regions (PARs) or presented paralogs in the genome. Thus, we ignored these five genes. Using the other 20 genes, we detected Y-chromosome gene expression as early as the 4-cell stage (TPM > 4), allowing embryo sexing at the time of genome activation (Figure 2A). Of the 23 activated embryos (disregarding the presumed parthenogenetic 16-cell embryo), we identified 12 male embryos: three male 8-cell embryos (J1, J2, D1), two male 16-cell embryos (S2, J2), four male morulae (S2, J1, J2, D2), and three male blastocysts (S2, J1, D2). Chromosomal expression levels are shown in Supplemental Figure S7.

As the location of the *XIST* gene had previously been reported in the EquCab2 genome [24], we used the Lift Genome Annotations tool from the UCSC Genome browser (https://genome.ucsc.edu/cgi-bin/hgLiftOver) to remap its location to the EquCab3 genome and evaluate its expression. In the EquCab3 genome version, *XIST* was located at the X:58,347,925-58,378,821 coordinates (reverse strand), which overlapped with the location of the human *XIST* gene in the non-horse refseq track. Female 8-cell embryos (as identified by the absence of Y-chromosome gene expression) exhibited significantly higher *XIST* expression than did 8-cell male embryos (18.0 ± 3.7 TPM in female vs. 3.2 ± 0.7 TPM in male; Padj <0.01). No significant differences were observed at the 16-cell stage, when male embryos increased the expression of *XIST* to a level that would be maintained across stages (TPM < 70) (Figure 2B). Based on the results of the Wald test, the female morulae presented higher expression levels (167.9 ± 9.1 TPM) than the male morulae (30.4 ± 22.1 TPM; $P < 0.01$), although the adjustment for multiple comparisons suggested this could be a false positive (Padj = 0.13). At the blastocyst stage, *XIST* expression from female embryos (388.4 ± 52.9 TPM) was high enough to clearly separate them from the male embryos (33.5 ± 10.8 TPM; FC > 10; Padj <0.01).

Differential expression between male and female embryos

After embryos were sexed, we evaluated differential gene expression between male and female embryos within each embryo stage (Supplemental File S4, Supplemental Figure S8). We found sex-related differentially expressed (S-DE) genes at all activated stages (Table 2). The 8-cell stage showed 226 S-DE genes with slightly marked downregulation in female embryos (69%), and 8% of the S-DE genes were located on the X chromosome. From the 16-cell stage onward, a high percentage of S-DE genes (~90%) were upregulated in female embryos, with most of the S-DE genes located on the X chromosome (40–62%).

The gene ontology (GO) analysis of the 25 S-DE genes from the 16-cell stage showed no significant enrichment for any GO terms, but according to the database, the genes upregulated in females were associated with poly(A) RNA binding, protein transport, protein localization, lysosomes and lytic vacuoles. The comparison between male and female morulae revealed 117 S-DE genes. Among these genes, we found critical factors for embryo development: *NOTCH2*, which encoded a factor involved in numerous cell-fate decisions, was upregulated in female embryos; while *CYP26A1*, which encoded for an enzyme that regulates cellular levels of retinoic acid, controlling important developmental processes such as sex differentiation, was

Table 2. Differential gene expression between male and female equine embryos produced by ICSI. As the list of GO terms was too large, only the less redundant and most informative terms were shown in this summary table. The full list of GO terms is available in Supplemental File S4

Embryo stage	Embryo sex	Upregulated genes	% of S-DE genes in X-chromosome	Enriched GO terms (Benjamini < 0.1)
8-cell	Female	71	8%	Cytosol
	Male	155		Cytoskeleton organization, cell morphogenesis, organelle organization, cellular component morphogenesis, regulation of cellular localization, enzyme binding, cytoskeletal protein binding, actin binding, kinase binding, protein binding, gene silencing, nuclear lumen, adherens junction organization, substrate adhesion-dependent cell spreading, ATP binding, embryo development, regulation of protein metabolic process, anchoring junction, DNA methylation involved in embryo development
16-cell	Female	23	40%	-
	Male	2		-
Morula	Female	106	62%	Chromosome organization, protein binding, covalent chromatin modification, nucleolus, methylation-dependent chromatin silencing, perinuclear region of cytoplasm, methyl-CpG binding, peptidyl-lysine modification, vesicle lumen, RNA metabolic process, positive regulation of proteolysis, protein stabilization, positive regulation of intracellular transport, RNA processing
	Male	11		-
Blastocyst	Female	65	58%	Extracellular exosome, extracellular vesicle, extracellular organelle, oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, secondary active transmembrane transporter activity, oxidoreductase activity, acting on CH-OH group of donors, active transmembrane transporter activity
	Male	1		-

downregulated in female embryos. In blastocysts, the number of S-DE genes decreased to 66, with 65 (98%) of these genes upregulated in female embryos and 38 (58%) of them on the X-chromosome. The *XIST* gene exhibited the highest upregulation in female embryos (FC = 11.3), which clearly distinguished *XIST* from the rest of the S-DE genes (FC < 7.2). The blastocyst S-DE genes included known factors such as *CDX2*, *PKD1*, *PLCD3* and *NSDHL*, involved in embryonic placenta development; and *WNT3A*, known for playing important roles in cell fate decisions during embryonic and fetal development. The only gene downregulated in female blastocysts was *ENSECAG00000030272*, which encoded for a lncRNA molecule and included one of the 5S rRNA genes at the 3′ end.

Biparental expression in the sex-chromosomes

Using the allele-specific data, we evaluated parental X-chromosome expression in female embryos to determine whether imprinted (parental-specific) X-inactivation occurred at the embryonic level. We detected biparental X-chromosome expression in the female embryos regardless of their developmental stage (Figure 3), similar to humans [38] and contrary to imprinted X-inactivation observed in mouse [39], cattle [21, 40], and marsupials [41]. In addition to demonstrating that both parental X-chromosomes were expressed at every stage in female embryos, these results corroborated embryo sexes inferred by Y-chromosome expression, as male embryos showed maternal X-chromosome expression, except for the pseudo-autosomal region (PAR) at the beginning of the X-chromosome and loci that could represent either sequencing errors or loci escaping inactivation.

Dosage compensation in the X-chromosome

To further evaluate X-chromosome gene expression in our embryo dataset, we assessed X-chromosome dosage compensation by calculating the relative expression of X-chromosome genes to autosomal genes (X:A ratio). As it has been demonstrated that noise can substantially affect X:A ratios, we analyzed genes with more than 4 TPM in each sample, and X-genes located outside pseudo-autosomal regions (selected by absence of paternal allele expression in XY embryos).

We observed incomplete dosage compensation ($0.5 < X:A < 1$) in all samples (male and female) except for the 16-cell and morula female embryos, which presented similar or even higher expression from the X chromosome(s) compared to autosomes (Figure 4). The increase in female X:A ratios leading up to this happened concurrently with genome activation, suggesting that the increase could be attributed to the activation of both X chromosomes. At the blastocyst stage, the female embryos showed a drop in their X:A ratios to levels like those before genome activation and similar to those of the male embryos. The significant accumulation of *XIST* RNA and the drop of X:A ratios in female blastocysts suggests X-inactivation might have started.

Discussion

This study represents the first critical evaluation of allele-specific gene expression during the first stages of development in the horse. In addition to increasing our knowledge of comparative embryo development, these data allowed us to validate early findings on the timing of EGA in equine embryos, as the signs of transcription

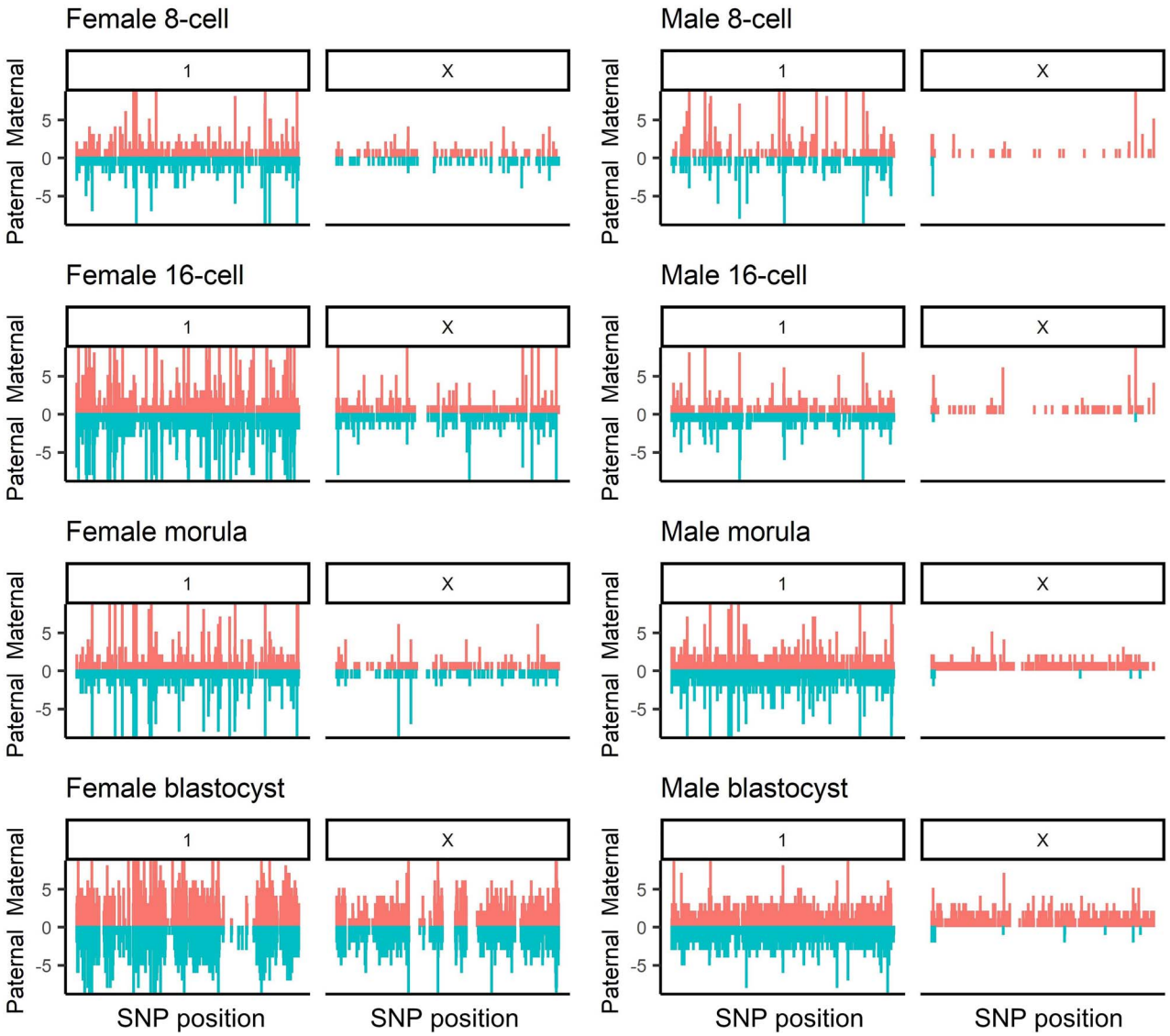

Figure 3. Parental expression (reads) in informative heterozygous loci from an autosome (chromosome 1) and the X-chromosome. The rest of the autosomes were left out to avoid redundance but presented similar patterns. Biparental expression was observed at chromosome 1 and the X-chromosome in female embryos from all stages, demonstrating the absence of imprinted X-inactivation at embryonic level reported in other species. Male embryos, on the other hand, showed mostly maternal allele expression in the X-chromosome, as expected, and a few loci expressing the paternal allele. Those loci located at the beginning of the X-chromosome (2 Mb) correspond to the pseudo-autosomal region, while the rest could represent either sequencing errors or regions escaping X-inactivation.

previously observed at the 5-cell [7] and 6-cell [8] stages occurred immediately after the increase of paternal gene expression that we detected at the 4-cell stage. However, it is worth mentioning that our results should not be considered representative of all horse embryos, as the embryos were produced in vitro. As a limitation, the use of a single stallion to reduce the genetic variability among embryos and to increase the probability of obtaining the same informative loci in multiple replicates, could have introduced a sire bias in the data. Another limitation was the abundance of duplicate reads in the dataset, which drastically reduced the number of informative loci with sufficient expression counts. Although this limitation prevented the analysis of allele-specific expression at most genes, including several control genes imprinted in horses and other species, the data were sufficient to capture parental contribution to global events such as genome activation and X-chromosome expression, as well as a few cases of allelic expression bias.

One such case was the *NANOGNB* gene, which showed paternal bias at the 16-cell, and potentially at the morula stage, although the expression counts in the second morula replicate were too low to conclude. This gene has been reported to play a role in early human development by downregulating genes expressed as a pulse during EGA [42]. Interestingly, paternal bias has also been observed in human morulae [11], suggesting evolutionary conservation and partially validating our findings in the horse. Another case of allelic bias was the *CUL2* gene, reported in horse placenta as well [12], although with the opposite parental origin. This difference suggested that the allelic bias we detected might have been caused by an eQTL, instead of a parental effect. This possibility applies to the eight cases

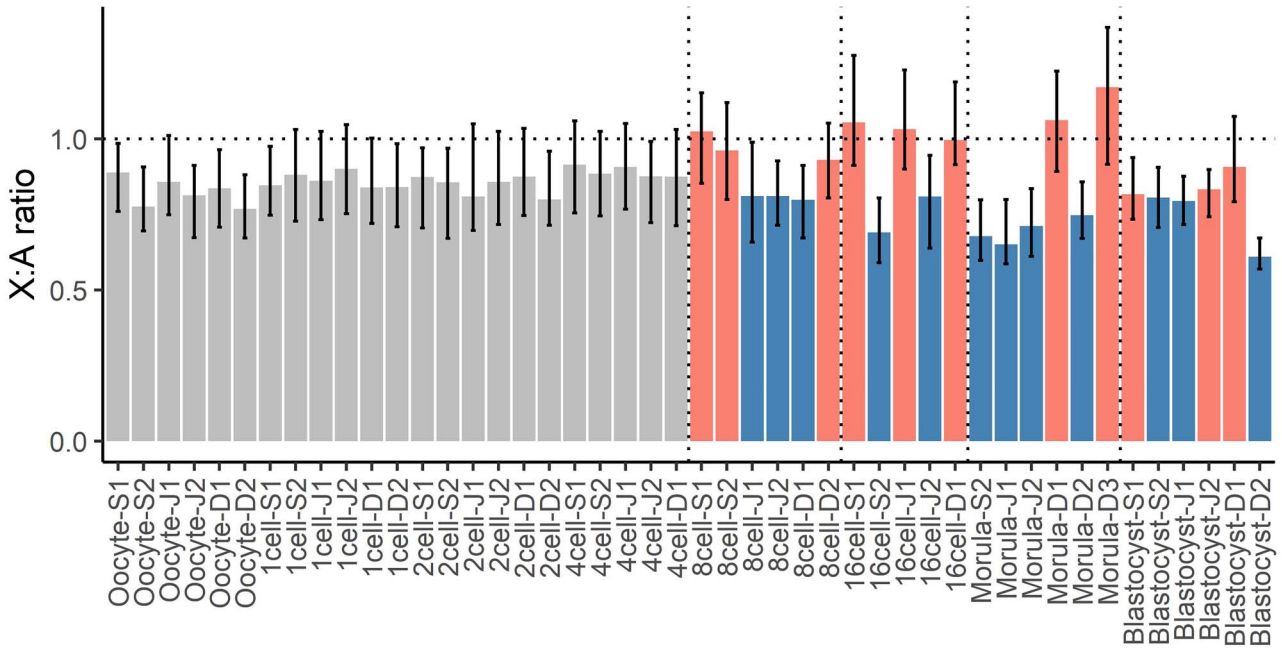

Figure 4. Expression of X-chromosome genes relative to autosomal genes (X:A ratios) in horse preimplantation embryos (blue = male; red = female). Scarce dosage compensation was observed in female embryos until the blastocyst stage. Bars represent 95% confidence intervals for the X:A ratios.

of allelic bias we found, as reciprocal parental crosses were not evaluated.

The importance of sexing embryos by detection of the Y-chromosome, levels of *XIST* expression, and heterozygous expression at the X-chromosome has recently been demonstrated [43]. We showed that horse embryos can be sexed right after genome activation at the 8-cell stage using all three approaches. As equine Y-chromosome assemblies were unavailable until recently, this work is, to the best of our knowledge, the first report of horse embryo sexing by Y-chromosome expression using NGS data. In horse preimplantation embryos, *XIST* is detected in both male and female embryos, as opposed to horse 35-day chorionic girdle and fetus samples, in which *XIST* was only detected in females [24]; and it is expressed immediately after genome activation at the 8-cell stage. In other species such cattle, XIST is expressed at the morula stage [44], even though EGA follows a similar timing. In addition, *XIST* expression is maintained at a relatively constant level in male embryos, whereas it significantly increases with development in female embryos, as reported in humans [45]. The lack of differences between male and female embryos at the 16-cell stage could have been caused by the high variation among embryos, although the existence of sex-specific regulatory mechanisms for *XIST* expression cannot be discarded.

Differences in gene expression between male and female embryos have previously been reported in other species, including mouse [46], cow [47, 48] and pig [49]. The list of S-DE genes given in these studies partially overlapped with ours, with some variation likely attributed to species-specific developmental processes or embryo production methods. As in other species, male equine embryos have been reported in equine commercial in vitro embryo production to develop more rapidly than do female embryos [50], thus the upregulation of the described genes in female equine embryos could be associated with slower embryonic development. The strong presence of the X-chromosome among S-DE genes seen in other species was conserved in the horse. Another interesting result of our analysis was the differential expression of genes associated with placental development at the blastocyst stage. In rats, sexual dimorphism has been reported in the structure and gene expression of E13 and E15 labyrinthine samples [51], although no differences have been reported in preimplantation embryos. Additional allele-specific expression studies in horse placental development could help clarify the importance of these S-DE genes.

Although we found no evidence of imprinted X-inactivation at the embryonic level, in contrast to mouse embryos, single cells could still express genes from one or both X-chromosomes. Thus, imprinted inactivation could still occur in extra-embryonic precursors starting to differentiate at the morula and blastocyst stages. To detect such phenomenon, additional studies based on single-cell RNA-seq are needed. Recent experiments using this technology in human preimplantation embryos have discarded the existence of imprinted X-inactivation in early human development [38, 52], establishing an important difference with respect to mouse development. In humans, dosage compensation is achieved by expression dampening of both X-chromosomes [38]; while in other mammalian species such as cattle, it is achieved by XCI [44]. Early cytogenetic analyses in equine conceptuses have reported the inactivation of one X-chromosome in trophoblastic cells at Day 7.5 [18], which is compatible with our findings and seems to indicate a similarity to the bovine model. Whether this mechanism is conserved in the horse is unclear, but we demonstrated that the equine mechanism does not involve imprinted inactivation.

The little dosage compensation observed at the 16-cell and morula stages seemed to constitute a horse-specific feature. For instance, human X-genes become gradually compensated after EGA until they reach almost complete compensation by Day 7 [38]. The reason for this difference is unclear; it could be associated with the pattern of *XIST* expression at the 16-cell and morula stages, but additional studies with larger sample sizes are necessary to

validate and investigate this interesting phenomenon. Independently of the mechanisms involved in the dosage compensation, we showed that autosomal and X-chromosome genes reach expression levels similar to those before EGA at the blastocyst stage. This timing was compatible with the cytogenetic report of XCI in the equine trophoblast at Day 7.5 [18], with XCI occurring in the embryonic disk a few days later (Day 11.5). However, single-cell RNA-Seq studies in peri-implantation embryos are necessary to answer this question.

Supplementary material

Supplementary material is available at *BIOLRE* online.

Author contribution statement

DEG performed the analyses and wrote the manuscript; PST produced the embryos and libraries; YHC produced the embryos; KH and PJR designed the experiments and supervised and edited the manuscript.

Acknowledgements

We would like to thank Deep Creek Farm, owned by Marsha Korose and Curtis Pierce, for donating semen of Norwegian Fjord Horse stallions to our project; the Texas A&M Theriogenology lab staff for their technical assistance on cryopreserving the semen samples; and Alma Islas-Trejo, Department of Animal Science, University of California, Davis, for her technical assistance during libraries preparation.

Data availability

Sequencing data are available at the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA688000.

Conflict of interest

Authors declare no conflicts of interest.

References

1. Aoki F, Hara KT, Schultz RM. Acquisition of transcriptional competence in the 1-cell mouse embryo: Requirement for recruitment of maternal mRNAs. *Mol Reprod Dev* 2003; **64**:270–274.
2. Flach G, Johnson MH, Braude PR, Taylor RA, Bolton VN. The transition from maternal to embryonic control in the 2-cell mouse embryo. *EMBO J* 1982; **1**:681–686.
3. Xue Z, Huang K, Cai C, Cai L, Jiang C, Feng Y, Liu Z, Zeng Q, Cheng L, Sun YE, Liu J, Horvath S et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 2013; **500**:593–597.
4. Graf A, Krebs S, Zakhartchenko V, Schwalb B, Blum H, Wolf E. Fine mapping of genome activation in bovine embryos by RNA sequencing. *PNAS* 2014; **111**:4139–4144.
5. Cao S, Han J, Wu J, Li Q, Liu S, Zhang W, Pei Y, Ruan X, Liu Z, Wang X, Lim B, Li N. Specific gene-regulation networks during the pre-implantation development of the pig embryo as revealed by deep sequencing. *BMC Genomics* 2014; **15**:4.
6. Chitwood JL, Burruel VR, Halstead MM, Meyers SA, Ross PJ. Transcriptome profiling of individual rhesus macaque oocytes and preimplantation embryos. *Biol Reprod* 2017; **97**:353–364.
7. Brinsko SP, Ball BA, Ignotz GG, Thomas PG, Currie WB, Ellington JE. Initiation of transcription and nucleologenesis in equine embryos. *Mol Reprod Dev* 1995; **42**:298–302.
8. Grøndahl C, Hyttel P. Nucleologenesis and ribonucleic acid synthesis in preimplantation equine embryos. *Biol Reprod* 1996; **55**:769–774.
9. McGrath J, Solter D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 1984; **37**:179–183.
10. Barton SC, Surani MA, Norris ML. Role of paternal and maternal genomes in mouse development. *Nature* 1984; **311**:374–376.
11. Leng L, Sun J, Huang J, Gong F, Yang L, Zhang S, Yuan X, Fang F, Xu X, Luo Y, Bolund L, Peters BA et al. Single-cell transcriptome analysis of uniparental embryos reveals parent-of-origin effects on human preimplantation development. *Cell Stem Cell* 2019; **25**: 697–712.e6.
12. Dini P, Kalbfleisch T, Uribe-Salazar JM, Carossino M, Ali HE-S, Loux SC, Esteller-Vico A, Norris JK, Anand L, Scoggin KE, Rodriguez Lopez CM, Breen J et al. Parental bias in expression and interaction of genes in the equine placenta. *Proc Natl Acad Sci U S A* 2021; **118**: e2006474118.
13. Heras S, Smits K, De Schauwer C, Van Soom A. Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during pronuclear development in equine zygotes produced by ICSI. *Epigenetics Chromatin* 2017; **10**:13.
14. Wang X, Miller DC, Harman R, Antczak DF, Clark AG. Paternally expressed genes predominate in the placenta. *Proc Natl Acad Sci U S A* 2013; **110**:10705–10710.
15. Stolakis V, Bertero MC. Molecular aspects of aneuploidy in preimplantation human embryos: A mini-review. *Reproductive Bio Medicine Online* 2019; **39**:e12–e13.
16. Gu P, Wang Y, Bisht KK, Wu L, Kukova L, Smith EM, Xiao Y, Bailey SM, Lei M, Nandakumar J, Chang S. Pot1 OB-fold mutations unleash telomere instability to initiate tumorigenesis. *Oncogene* 2017; **36**:1939–1951.
17. Rappold GA. The pseudoautosomal regions of the human sex chromosomes. *Hum Genet* 1993; **92**:315–324.
18. Romagnano A, Richer CL, King WA, Betteridge KJ. Analysis of X-chromosome inactivation in horse embryos. *J Reprod Fertil Suppl* 1987; **35**:353–361.
19. Oikawa M, Inoue K, Shiura H, Matoba S, Kamimura S, Hirose M, Mekada K, Yoshiki A, Tanaka S, Abe K, Ishino F, Ogura A. Understanding the X chromosome inactivation cycle in mice: A comprehensive view provided by nuclear transfer. *Epigenetics* 2014; **9**:204–211.
20. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* 2004; **303**:644–649.
21. Ferreira AR, Machado GM, Diesel TO, Carvalho JO, Rumpf R, Melo EO, Dode MAN, Franco MM. Allele-specific expression of the MAOA gene and X chromosome inactivation in in vitro produced bovine embryos. *Mol Reprod Dev* 2010; **77**:615–621.
22. Richardson BJ, Czuppon AB, Sharman GB. Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. *Nat New Biol* 1971; **230**:154–155.
23. Sharman GB, Late DNA. Replication in the paternally derived X chromosome of female kangaroos. *Nature* 1971; **230**:231–232.
24. Wang X, Miller DC, Clark AG, Antczak DF. Random X inactivation in the mule and horse placenta. *Genome Res* 2012; **22**:1855–1863.
25. McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guérin G, Hasegawa T, Hill EW, Leeb T, Lindgren G et al. A high density SNP array for the domestic horse and extant Perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet* 2012; **8**:e1002451.
26. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011; **43**:491–498.
27. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 2009; **25**:1754–1760.
28. Janečka JE, Davis BW, Ghosh S, Paria N, Das PJ, Orlando L, Schubert M, Nielsen MK, Stout TAE, Brashear W, Li G, Johnson CD et al. Horse Y chromosome assembly displays unique evolutionary features and putative stallion fertility genes. *Nat Commun* 2018; **9**:2945.

29. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**:15–21.
30. Liao Y, Smyth GK, Shi W. Feature counts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014; **30**:923–930.
31. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009; **4**:1184–1191.
32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**:550.
33. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**:44–57.
34. Walter W, Sánchez-Cabo F, Ricote M. GOplot: An R package for visually combining expression data with functional analysis. *Bioinformatics* 2015; **31**:2912–2914.
35. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011; **27**:2987.
36. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R et al. The variant call format and VCFtools. *Bioinformatics* 2011; **27**:2156–2158.
37. Dirks RAM, van Mierlo G, Kerstens HHD, Bernardo AS, Kobolák J, Bock I, Maruotti J, Pedersen RA, Dinnyés A, Huynen MA, Jouneau A, Marks H. Allele-specific RNA-seq expression profiling of imprinted genes in mouse isogenic pluripotent states. *Epigenetics Chromatin* 2019; **12**:14.
38. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, 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.
39. Heard E, Disteche CM. Dosage compensation in mammals: Fine-tuning the expression of the X chromosome. *Genes Dev* 2006; **20**:1848–1867.
40. Duan J (Ellie), Shi W, Jue NK, Jiang Z, Kuo L, O'Neill R, Wolf E, Dong H, Zheng X, Chen J, Tian X (Cindy). Dosage compensation of the X chromosomes in bovine germline, early embryos, and somatic tissues. *Genome Biol Evol* 2018; **11**:242–252.
41. Al Nadaf S, Waters PD, Koina E, Deakin JE, Jordan KS, Graves JA. Activity map of the tammar X chromosome shows that marsupial X inactivation is incomplete and escape is stochastic. *Genome Biol* 2010; **11**:R122.
42. Dunwell TL, Holland PWH. A sister of NANOG regulates genes expressed in pre-implantation human development. *Open Biol* 2017; **7**:170027.
43. Midic U, Vande Voort CA, Latham KE. Determination of single embryo sex in Macaca mulatta and Mus musculus RNA-Seq transcriptome profiles. *Physiol Genomics* 2018; **50**:628–635.
44. Yu B, van Tol HTA, Stout TAE, Roelen BAJ. Initiation of X chromosome inactivation during bovine embryo development. *Cell* 2020; **9**:E1016.
45. van den Berg IM, Laven JSE, Stevens M, Jonkers I, Galjaard R-J, Gribnau J, Hikke van Doorninck J. X chromosome inactivation is initiated in human preimplantation embryos. *Am J Hum Genet* 2009; **84**:771–779.
46. Lowe R, Gemma C, Rakyan VK, Holland ML. Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development. *BMC Genomics* 2015; **16**:295.
47. Bermejo-Alvarez P, Rizos D, Rath D, Lonergan P, Gutierrez-Adan A. Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. *Proc Natl Acad Sci U S A* 2010; **107**:3394–3399.
48. Chitwood JL, Rincon G, Kaiser GG, Medrano JF, Ross PJ. RNA-seq analysis of single bovine blastocysts. *BMC Genomics* 2013; **14**:350.
49. 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**:679.
50. Claes A, Cuervo-Arango J, Colleoni S, Lazzari G, Galli C, Stout TA. Speed of in vitro embryo development affects the likelihood of foaling and the foal sex ratio. *Reprod Fertil Dev* 2020; **32**:468–473.
51. Kalisch-Smith JI, Simmons DG, Pantaleon M, Moritz KM. Sex differences in rat placental development: From pre-implantation to late gestation. *Biol Sex Differ* 2017; **8**:17.
52. Moreira de Mello JC, Fernandes GR, Vibranovski MD, Pereira LV. Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. *Sci Rep* 2017; **7**:10794.