Parental bias in expression and interaction of genes in the equine placenta

Pouya Dini[a,b,1], Theodore Kalbfleisch[c,2], José M. Uribe-Salazar[d], Mariano Carossino[a,3], Hossam El-Sheikh Ali[a,e], Shavahn C. Loux[a], Alejandro Esteller-Vico[a,4], Jamie K. Norris[a], Lakshay Anand[f], Kirsten E. Scoggin[a], Carlos M. Rodriguez Lopez[f], James Breen[g], Ernest Bailey[a], Peter Daels[b], and Barry A. Ball[a,5]

[a]Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington, KY 40503; [b]Department of Veterinary Medical Imaging and Small Animal Orthopaedics, Faculty of Veterinary Medicine, Ghent University, Merelbeke 9820, Belgium; [c]Department of Biochemistry and Molecular Genetics, University of Louisville, Louisville, KY 40202; [d]Department of Biochemistry and Molecular Medicine, Genome Center, Medical Investigation of Neurodevelopmental Disorders Institute, University of California, Davis, CA 95616; [e]Theriogenology Department, Faculty of Veterinary Medicine, University of Mansoura, 35516, Egypt; [f]Environmental Epigenetics and Genetics Group, Department of Horticulture, University of Kentucky, Lexington, KY 40546; and [g]South Australian Health and Medical Research Institute, Adelaide, SA 5001, Australia

Edited by George E. Seidel, Colorado State University, Fort Collins, CO, and approved February 24, 2021 (received for review April 6, 2020)

Most autosomal genes in the placenta show a biallelic expression pattern. However, some genes exhibit allele-specific transcription depending on the parental origin of the chromosomes on which the copy of the gene resides. Parentally expressed genes are involved in the reciprocal interaction between maternal and paternal genes, coordinating the allocation of resources between fetus and mother. One of the main challenges of studying parental-specific allelic expression (allele-specific expression [ASE]) in the placenta is the maternal cellular remnant at the fetomaternal interface. Horses (*Equus caballus*) have an epitheliochorial placenta in which both the endometrial epithelium and the epithelium of the chorionic villi are juxtaposed with minimal extension into the uterine mucosa, yet there is no information available on the allelic gene expression of equine chorioallantois (CA). In the current study, we present a dataset of 1,336 genes showing ASE in the equine CA (https://pouya-dini.github.io/equine-gene-db/) along with a workflow for analyzing ASE genes. We further identified 254 potentially imprinted genes among the parentally expressed genes in the equine CA and evaluated the expression pattern of these genes throughout gestation. Our gene ontology analysis implies that maternally expressed genes tend to decrease the length of gestation, while paternally expressed genes extend the length of gestation. This study provides fundamental information regarding parental gene expression during equine pregnancy, a species with a negligible amount of maternal cellular remnant in its placenta. This information will provide the basis for a better understanding of the role of parental gene expression in the placenta during gestation.

allele-specific expression | parental gene expression | monoallelic gene expression | placenta | equine

The placenta is the fetomaternal interface, which is essential for fetal growth and survival. In eutherian mammals, the placenta plays a pivotal role in producing hormones, transporting nutrients and byproducts between the fetal and maternal circulation, and protecting the fetus from the maternal immune system (1, 2). A prerequisite for the efficient function of the placenta is dynamic gene expression throughout gestation, mediating the rapid physiological changes in this tissue during pregnancy (3, 4). Efforts toward assigning a parent of origin to the genes critically involved in this process has remained an active research topic during the last decade.

In mammalian cells, biallelic expression of autosomal genes constitutes the most common form of gene expression, by which both alleles (paternally and maternally derived) are simultaneously transcribed and therefore equally contribute to the gene expression profile. In contrast, there are groups of genes that exhibit monoallelic expression where allele-specific transcription is dependent on the parental origin of the chromosomes (5–7). Interestingly, this parent-specific transcription significantly contributes to the placental development and function (8). It has been shown that genes with parental bias in their expression are critical for normal embryonic growth, including the development of organs, the modulation of placental surface area, and vascular branching density, among numerous other functions (9–13). Consequently, alterations in the expression of parentally expressed genes have been linked to several developmental disorders of the

Significance

The placenta is the fetomaternal interface, which is essential for fetal growth and survival, playing a central role in the health of both the fetus and its mother. The dynamic gene expression during pregnancy dramatically contributes to the correct functionality of this temporary tissue. The epitheliochorial placenta of the horse is a valuable resource to understand parent-of-origin expression due to minimal bias associated with remnants of maternal tissue compared to other eutherian mammals. Here, we identified genes whose transcription is biased to either the paternal or maternal chromosome in the equine placenta. Overall, this study contributes to a better understanding of regulatory processes in placental function, evolution, and disease, using horses as a model for eutherian mammals' placenta.

Author contributions: P. Dini designed the study; T.K., M.C., H.E.-S.A., S.C.L., A.E.-V., and E.B., provided help in the study design; J.M.U.-S. and L.A. designed the webpage of the dataset; P. Dini performed all experimental procedures including animal work, bioinformatic, and statistical analyses; J.M.U.-S., J.K.N., and L.A. assisted with the bioinformatic analysis; S.C.L. and A.E.-V. assisted with the animal work; P. Dini prepared figures and tables; M.C., H.E.-S.A., J.K.N., and L.A. provided help with figure preparation; T.K. wrote the in-house java program; M.C., C.M.R.L., and J.B. provided help in the data curation; K.E.S. contributed to sample preparation; P. Dini wrote the manuscript; T.K., J.M.U.-S., M.C., H.E.-S.A., S.C.L., A.E.-V., K.E.S., C.M.R.L., J.B., and E.B. revised and edited versions of the manuscript; P. Daels oversaw the entire project, including the acquisition of funding, experimental design, and review of the manuscript; and B.A.B. oversaw the entire project, including the acquisition of funding, animal work, statistics, and manuscript writing and submission.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]Present address: Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616.

[2]Present address: Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington, KY 40503.

[3]Present address: Louisiana Animal Disease Diagnostic Laboratory, Department of Pathobiological Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA 70803.

[4]Present address: Department of Biomedical and Diagnostic Sciences, College of Veterinary Medicine, University of Tennessee, Knoxville, TN 37996.

[5]To whom correspondence may be addressed. Email: b.a.ball@uky.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2006474118/-/DCSupplemental.

Published April 14, 2021.

placenta, miscarriage, preeclampsia, and intrauterine growth restriction in women, mice, and domestic animals (14–22).

It has been suggested that parental gene expression, including imprinted genes, arose during the course of mammalian evolution to facilitate the interaction between parent and fetus (23, 24). The reciprocal interaction between maternal and paternal gene expression coordinates the allocation of resources between the fetus and the mother (25, 26). Overall, it has been proposed that the paternally expressed genes tend to maximize the resources that are received by the offspring (enhancer of fetal growth), whereas the maternally expressed genes tend to reduce the distribution of these resources to the offspring in an attempt to modulate the maternal energy supply during pregnancy, a phenomenon known as parental conflict theory (9, 16, 25–27). It is noteworthy that parental genes are not limited to protein-coding genes since the expression of specific noncoding RNAs (ncRNAs) can also be parentally expressed (26). Several functions have been allocated to ncRNAs through their RNA–RNA interactions in many cell types (28–34). The majority of these ncRNAs are either located in yet unannotated regions of the genome or are located within the opposite strand of the expressed protein-coding gene (35). Recently, it has been revealed that this type of antisense transcription is common in mammals and other eukaryotes, where the antisense transcript regulates the expression of the gene located on the sense strand either transcriptionally or at the posttranscriptional level (30, 36–45).

Advances in high-throughput sequencing technologies have provided a unique opportunity to identify the parentally expressed loci, including imprinted genes, in the murine and human placenta (46–49). However, one of the challenges of studying these genes is associated with their hemochorial placentation and associated mixed maternal and fetal cellular components at the fetomaternal interface (50). This results in the overestimation of maternally expressed genes in the placenta (49). This obstacle can be overcome by studying parental expression of genes in the placenta of species with epitheliochorial and a minimally invasive type of placentation, such as the equine placenta. The horse has an epitheliochorial placenta in which both the endometrial epithelium and the epithelium of the chorionic villi are juxtaposed with minimal extension into the uterine mucosa. Therefore, the fetal placenta (chorioallantoic membrane) can be separated from the endometrium with negligible cellular mixture from maternal components (51). Currently, there is no information available for the allele-specific gene expression in the equine placenta except the study by Wang et al., in which the authors identified parental gene expression in the equine chorionic girdle collected from horse–donkey hybrids (52). Although this was a well-executed study, the functions and the characteristics of the chorionic girdle and chorioallantoic membranes are significantly different. The chorionic girdle is a transient circular band with a distinctly high paternal influence that invades the endometrium and forms the endometrial cups during the first trimester of equine pregnancy (53–55). Currently, there is no information available on the parental gene expression of equine chorioallantois (CA). We hypothesized that there are several transcripts (protein-coding and nonprotein coding) with parental bias in their expression in equine CA. Therefore, the aim of our study was to identify loci with parental bias in their expression in equine CA during mid-gestation and to provide a dataset of these genes along with their expression pattern throughout gestation.

Results

Identification of Allele-Specific Expressed Genes in Equine CA by High-Throughput Sequencing. To identify allele-specific expression (ASE) loci in equine placenta, we performed RNA-sequencing (RNA-seq) (stranded and paired-end; see *Materials and Methods* and *SI Appendix*, Fig. S1*A*) on CA (n = 6) in conjunction with genomic whole genome shotgun sequencing (WGS, mean ± SD coverage of 28X ± 3X) on CA, along with both the dam and sire, generating six sets of parent–offspring trios (parent–placenta trios). RNA-seq reads from each CA sample were phased according to the strand of the transcribed RNA [forward and reverse strands (45)], and expressed alleles at each location were compared to the respective sire and dam genomic DNA (gDNA). Using the RNA-seq data, we identified single-nucleotide polymorphisms (SNPs) to distinguish the parent of origin of the transcripts. The expressed alleles (carrying SNPs at heterozygous sites) which could only originate from one of the parents (both parent and offspring homozygous at the respective site) were considered informative and selected for further analysis. From a total of 2,853,000 SNPs detected, 570,334 (~19%) informative SNPs (excluding mitochondrial SNPs) were identified in the RNA-seq data from the CA. To confirm that the detected alleles were not due to artifacts from the sequencing, the presence of each informative SNP was confirmed in the gDNA of the corresponding CA (via WGS). Out of 570,334 informative SNPs detected in the RNA-seq, 549,593 (96%) SNPs were confirmed via WGS of the CA (*SI Appendix*, Table S1). The resulting SNP density was similar among the CA samples, with a mean ± SD of 0.08 ± 0.01 informative SNPs/kb [*Equus caballus* genome size of ~2.6×10^6 (56, 57)].

Initially, informative SNPs from all the parent–placenta trios that were located on annotated regions of genome were combined (based on the transcribed RNA strand; forward and reverse) to obtain gene-level representation. A total of 435,157 informative SNPs were identified in the annotated loci, corresponding to 18,446 genes out of 33,460 annotated genes (~55%) existing in the *Equus caballus* reference transcriptome (EquCab 3.0, National Center for Biotechnology Information [NCBI]) (Dataset S1). We set an arbitrary threshold of a minimum of six SNPs per gene, thus removing genes which had less than six informative SNPs (a total of 4,993 genes were removed). On average, 32.1 ± 38.9 (mean ± SD) SNPs were included per gene (combining all samples). Next, the distribution of paternal and maternal SNPs per gene was calculated. Several cutoff points have been used in previous studies to distinguish between allele-specific and biallelic-expressed genes; the cutoff values vary from 60% paternal or maternal expression to as restrictive as 100% expression originating from one of the parents (52, 58–61). In the present study, we used the cutoff of maternal expression (Mat) ≥75% and paternal expression (Pat) ≤25% for maternally expressed candidates and Mat ≤25% and Pat ≥75% for paternally expressed genes, respectively (Fig. 1*A*), to identify the significant allele-specific expressed genes (ASEGs) (8). Out of 13,453 genes with informative SNPs, 2,022 (~15%) exhibited an ASE pattern [≥75% paternal or maternal, false discovery rate (FDR) < 0.05, P value = 0.01], with 57% (1,157 genes) being maternally expressed and 43% (865 genes) paternally expressed. In order to increase the confidence level, we further separated these ASEGs into four groups based on the parental bias in their expression: Group A, 100% (n = 269 Mat and 163 Pat genes); Group B, 100 > to > 90% (n = 277 Mat and 139 Pat genes); Group C, 90 ≥ to > 80% (n = 393 Mat and 340 Pat genes); and Group D, 80 ≥ to ≥ 75% (n = 218 Mat and 223 Pat genes) (Fig. 1*B* and Dataset S2).

Confirmation of Parental Bias in the Expression of ASEGs for Each Individual Parent–Offspring Trio. As the initial approach of summing up the information (informative SNPs) from all the trios might have masked and biased the results in identifying the ASEGs, we analyzed the predicted ASEGs in each individual trio. To maintain their ASEGs status, these needed to demonstrate the same parental bias in their expression within each trio. Subsequently, the list of ASEGs was narrowed to the genes that show parental bias in their expression in at least three parent–placenta trios. A total of 190 genes (out of 2,022) were removed

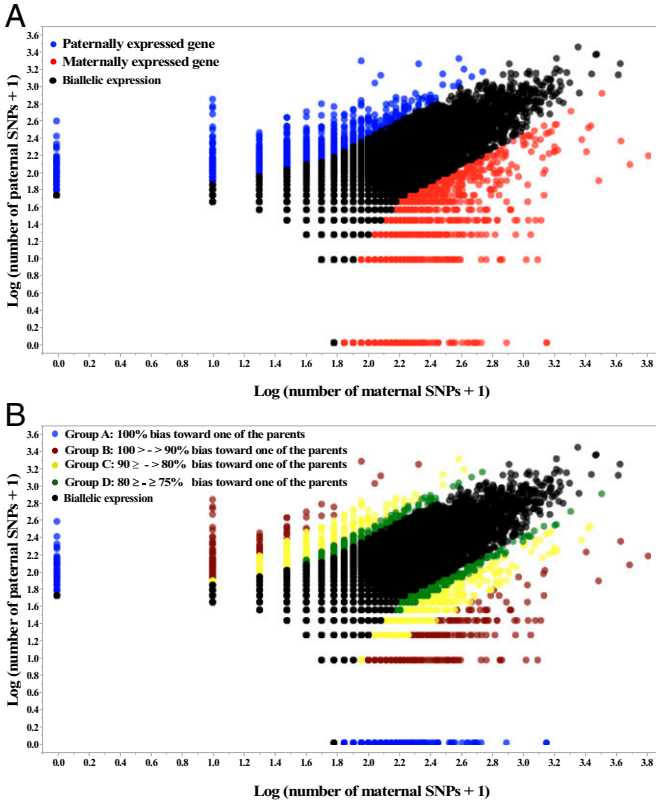

Fig. 1. (*A*) The number of informative SNPs in parentally expressed and biallelic expressed genes (each dot represents a gene). A binomial test was used to detect statistically significant deviations between number of maternal and paternal SNPs for each gene. Multiple testing-corrected *P* values were calculated by the false discovery rate method (Benjamini and Hochberg FDR < 0.05). To identify the significant ASEGs, we used a cutoff of maternal expression (Mat) >75% and paternal expression (Pat) <25% for maternally expressed candidates (red dots) and Mat <25% and Pat >75% for paternally expressed ones (blue dots). Black dots represent the biallelically expressed genes. The scatter plot of the ASEGs was generated using Package "tau-charts" in R. The log (number of SNPs + 1) is represented in the *x*- and *y*-axes. Gene names can be found in the interactive scatter plot at https://rpubs.com/pouyadini/545723. (*B*) Groups of ASEGs based on four different cutoff points for detection of parental bias (each dot represents a gene). ASEGs were grouped and presented in four groups based on the parental bias in their expression (Group A: 100% [n = 269 Mat and 163 Pat genes]; Group B: 100 > to ≥ 90% [n = 277 Mat and 139 Pat genes]; Group C: 90 > to ≥ 80% [n = 393 Mat and 340 Pat genes]; and Group D: 80 > to ≥ 75% [n = 218 Mat and 233 Pat]). Gene names can be found in the interactive scatter plot at https://rpubs.com/pouyadini/545721. Dots are colored by grouping rather than by parent of origin.

from the initial list of ASEGs (*SI Appendix*, Table S2). Genes that were not sufficiently expressed (average expression of fragments per kilobase of transcript per million fragments mapped [FPKM] < 1 among the six CA samples) were also removed from the list of ASEGs (n = 497 genes were removed). Fig. 2*A* represents the number of ASEGs across the different trios. A total of 37 genes (out of 202 genes in Group A) had informative SNPs in at least five of the trios and had 100% bias in their SNPs toward one of the parents, providing the highest confidence in the prediction of their allelic expression (Table 1). Among the 1,336 ASEGs, 1,111 expressed genes were autosomal, and 225 genes were located on the X chromosome (Dataset S3). The male and female CA were treated as a biological replicate in this study; therefore, we removed the chromosome X from the downstream analysis. The distribution of ASEGs (526 Mat and 585 Pat genes) across the equine chromosomes is presented in Fig. 2*B*.

Among the 1,336 ASEGs, 1,291 were coding and 45 were noncoding genes (Dataset S3). Lastly, we performed functional annotations for the protein-coding ASEGs. Gene ontology (GO) analysis of maternally and paternally expressed genes demonstrated that paternally expressed genes were mainly involved in metabolic and biosynthesis processes of proteins, macromolecules, and organic compounds, while maternally expressed genes were involved in localization, macromolecule and protein modifications, positive regulation of cell death, and apoptotic process (Dataset S4 and Fig. 3).

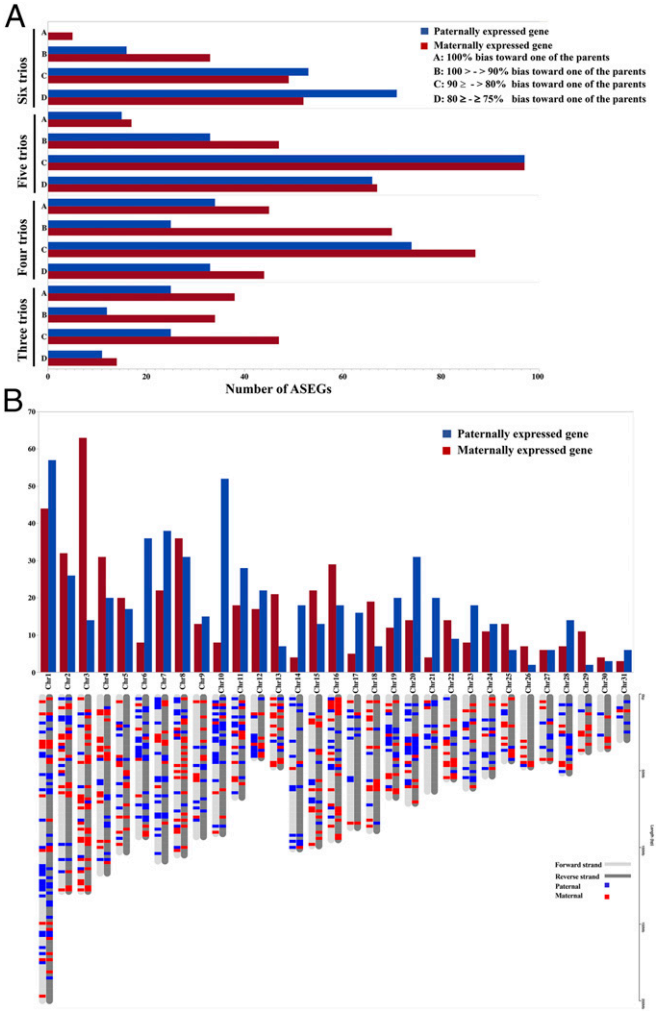

Fig. 2. (*A*) ASEGs in the individual parent–placenta trios. Combining the informative SNPs from all the trios may bias the results in identifying the ASEGs. Therefore, we analyzed the predicted ASEGs in each individual trio. The number of ASEGs in the parent–placenta trios are classified based on the number of trios that showed the informative SNPs. Red and blue bars indicate maternally and paternally expressed genes, respectively. (*B*) The distribution of ASEGs across equine chromosomes (*Top*). A total of 1,111 genes were located on autosomal chromosomes. ECA1 had the largest number of ASEGs among the autosomal chromosomes followed by ECA10, ECA3, and ECA7. A binomial test showed that there were biases in the proportion of maternal and paternal genes (toward one of the parents) in chromosomes on ECA3 and ECA10. (*Bottom*) Ideogram with the genomic map of the ASEGs. Distribution of paternally and maternally expressed genes are demonstrated based on the strand of their transcription. The ideogram was generated using Package "ChromoMap" in R (62). Gene names can be found in the interactive ideogram at https://rpubs.com/pouyadini/594229.

GENETICS

Table 1. ASEGs with informative SNPs in at least five out of the six studied trios and 100% bias in the SNPs toward one of the parents

Gene name	Chromosome	Number of SNPs	Number of trios	Parental expression	Binomial distribution FDR	Transcription direction
SLC16A12	Chr1	11	5	Maternal	0.005	Reverse
ANKRD1	Chr1	10	6	Maternal	0.009	Reverse
SNRPN	Chr1	47	5	Paternal	<0.001	Reverse
COQ9	Chr3	10	5	Paternal	0.009	Reverse
PLAC8	Chr3	7	5	Maternal	0.039	Reverse
MON1B	Chr3	8	5	Maternal	0.024	Reverse
FKBP14	Chr4	8	5	Maternal	0.025	Forward
DGKA	Chr6	9	5	Paternal	0.015	Reverse
STK25	Chr6	9	5	Paternal	0.015	Forward
RAB38	Chr7	7	5	Paternal	0.039	Reverse
IL10RA	Chr7	13	5	Maternal	0.002	Reverse
HIP1R	Chr8	14	5	Maternal	0.001	Reverse
VPS37B	Chr8	11	6	Maternal	0.005	Forward
DCAF13	Chr9	12	5	Paternal	0.003	Reverse
LOC111775242	Chr10	15	5	Paternal	0.001	Forward
ZNF544	Chr10	7	5	Paternal	0.039	Reverse
ALOXE3	Chr11	11	5	Maternal	0.005	Forward
B4GALNT2	Chr11	15	6	Maternal	0.001	Reverse
FGF19	Chr12	9	5	Paternal	0.015	Forward
RNF145	Chr14	12	5	Paternal	0.003	Reverse
CLK1	Chr18	8	5	Maternal	0.025	Forward
PLCXD2	Chr19	16	5	Maternal	<0.001	Forward
EIF2B5	Chr19	12	5	Paternal	0.003	Reverse
FITM2	Chr22	10	5	Maternal	0.009	Forward
LOC111770146	Chr23	8	5	Paternal	0.024	Reverse
LOC102149687	Chr23	9	5	Paternal	0.015	Forward
NDUFB1	Chr24	10	5	Paternal	0.009	Forward
DRAM1	Chr28	16	5	Paternal	<0.001	Reverse
NLGN4X	ChrX	139	6	Maternal	<0.001	Forward
USP9X	ChrX	33	5	Maternal	<0.001	Forward
WDR45	ChrX	22	5	Maternal	<0.001	Forward
TMEM185A	ChrX	23	5	Maternal	<0.001	Forward
ARHGEF9	ChrX	13	5	Maternal	0.002	Forward
SLC9A7	ChrX	34	5	Maternal	<0.001	Forward
RLIM	ChrX	7	5	Maternal	0.04	Forward
LAS1L	ChrX	17	6	Maternal	<0.001	Forward
CUL4B	ChrX	27	5	Maternal	<0.001	Forward

Mitochondrial SNPs with Maternal Origin. To further validate our methodology for detection of ASEGs, parental expression bias of informative SNPs in mitochondrial RNA was analyzed. A total of 930 SNPs were detected in the mitochondrial RNA with 342 SNPs considered as informative (following the previously described methodology). Out of these 342 informative SNPs, three SNPs showed paternal origin and 339 were of maternal origin, as expected, since mitochondrial DNA is inherited exclusively from the mother (Dataset S5). A total of 327 informative SNPs were further confirmed in the WGS dataset, eliminating the three paternal SNPs. The three nonconfirmed SNPs were all detected in the same trio.

Methylation Status of ASEGs in the Equine Placenta. To investigate the origin of the allelic expression of the identified ASEGs, DNA methylation of these genes was analyzed in the CA samples (n = 6) using reduced representation bisulfite sequencing (RRBS). After mapping the obtained reads, the methylation level for each methylated cytosine was calculated using a binomial distribution (see *Materials and Methods*). Methylated cytosines with a sequencing depth ≥5 and a corrected P value from the binomial test ≤0.01 were used for further analysis (64, 65). In total, 64,627 significantly methylated cytosines were identified in the placenta samples (Dataset S6). These methylated cytosines were annotated to the equine reference transcriptome, corresponding to 2,894 autosomal genes (Fig. 4*A*). Next, the list of autosomal ASEGs (n = 1,111) was compared to the list of the genes with identified methylated cytosines. A total of 254 genes, including *IGF1R*, *IGF2*, *IGF2R*, *RTL1*, *DLK1*, and *INSR*, exhibited both allele-specific expression and significant methylated sites, suggesting they could potentially be imprinted (Fig. 4 *A* and *B* and Dataset S7). A total of 21 genes belonged to Group A (7 Mat and 14 Pat; Fig. 4*C*), 28 to Group B (11 Mat and 17 Pat; Fig. 4*C*), 118 to Group C (65 Mat and 53 Pat), and 87 to Group D (46 Mat and 41 Pat; Dataset S7).

We compared the list of previously reported imprinted genes in chorionic girdle of equine hybrids (52) to our potentially imprinted gene list. Out of 24 reported imprinted genes, 4 (*IGF2*, *IGF2R*, *DLK1*, and *INSR*) were identified in our list (*SI Appendix*, Table S3). Other imprinted genes such as *H19*, *PHLDA2*, *MEST*, *NDN*, and *LY6G6C* were excluded from our list due to a low number of SNPs (*SI Appendix*, Table S3). Other genes (*PEG10*, *PEG3*, *NAP1L4*, and *SNRPN*, among others) were identified as ASEGs in our dataset; however, we were not able to demonstrate significant methylated sites in these genes. We detected no discrepancy between the parental expression of the SNPs in the two datasets. Wang et al. (52) reported *DIRAS3* as a paternally imprinted gene based on one significant SNP from their RNA-seq data. In our dataset, we found three paternally expressed SNPs in the sense and four SNPs (two maternally and two paternally expressed SNPs) in the antisense strand. However, since the dataset used by Wang et al. (52) was unstranded, it is unclear which strand of DNA was the origin of the paternal SNP identified by them. Additionally, some of the

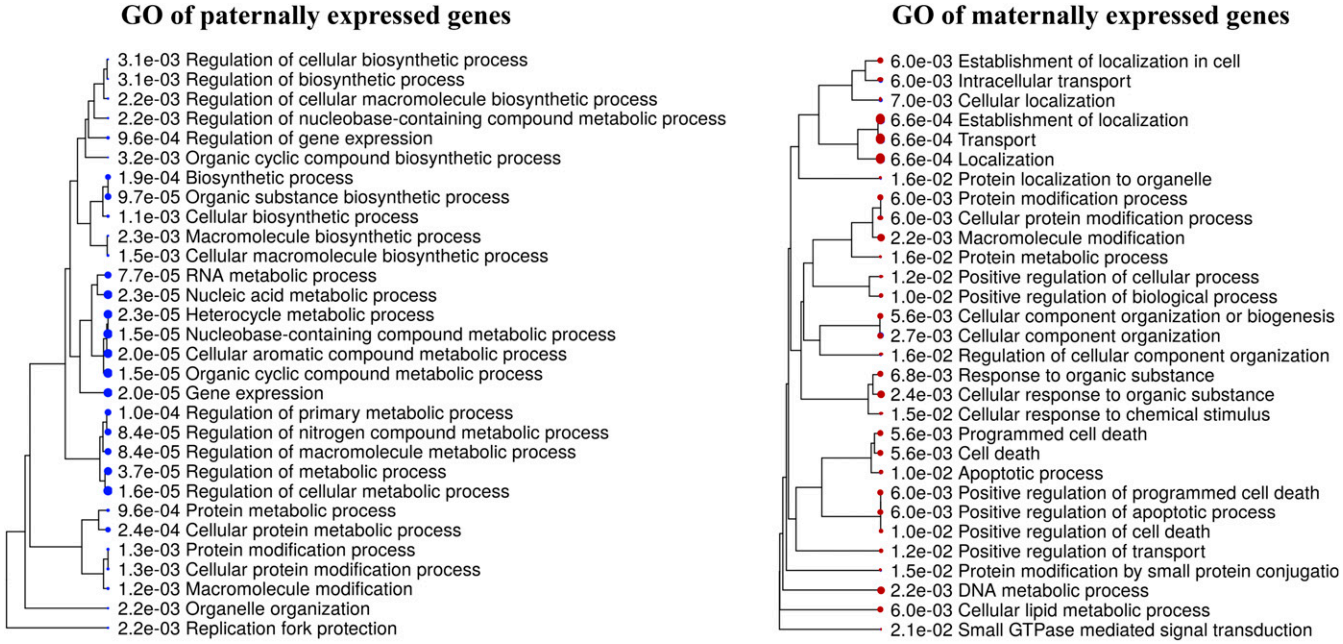

Fig. 3. Hierarchical clustering tree of predicted GO of paternally and maternally expressed genes. The size of the nodes corresponds to the number of genes involved in each process (63).

imprinted genes in the chorionic girdle, such as *INS*, *PAR-SN*, and *D7ERTD715E* (52), were not expressed in CA. This again reflects the differences between the nature of chorionic girdle and CA.

GO analysis of maternally and paternally expressed genes with significant methylated sites suggested that biological pathways involved in positive regulation of cell death, positive regulation of programmed cell death, and positive regulation of apoptotic process are enriched in maternally expressed genes (Fig. 4*D* and Dataset S8). On the other hand, pathways related to the negative regulation of cell cycle checkpoints, negative regulation on DNA damage checkpoint, establishment of microtubule cytoskeleton polarity, and regulation of protein kinase B signaling were enriched in paternally expressed genes (Fig. 4*D* and Dataset S8).

Expression Patterns of Potentially Imprinted Genes throughout the Equine Gestation. We evaluated the expression patterns of the genes we identified as potentially being imprinted genes throughout the equine pregnancy (45 d, 4 mo, 6 mo, and 10 mo; $n = 4$ for each time point) using a previously generated paired-end, RNA-seq dataset from CA (Fig. 5 and Dataset S9) (66). To evaluate the kinetics of these potentially imprinted genes throughout gestation, differentially expressed genes (DEGs) during the pregnancy were identified (FDR < 0.05). A total of 51 maternally and 58 paternally expressed genes showed differential expression across gestational time points (Fig. 6 and Dataset S10). We further classified the GO of DEGs. Maternal DEGs were involved in transmembrane transporter activity, cyclic guanosine monophosphate-dependent protein kinase (cGMP-PKG) signaling, semaphoring-plexin signaling, and fluid shear stress. Paternal DEGs were involved in insulin-related signaling, negative regulation of DNA damage checkpoint, negative regulation of cell cycle checkpoint, catalytic activity, and transmembrane transport (Dataset S11).

Next, we grouped the DEGs into three clusters using the k-means clustering method (Fig. 6). The expression of two clusters consisting of 24 maternally expressed genes increased during pregnancy. There was also a cluster of maternally expressed genes in which the genes had higher expression at the beginning of the pregnancy and then declined toward the end of pregnancy. Within the paternally expressed genes, there was one cluster ($n = 25$ genes) which showed an increasing trend, and two clusters ($n = 32$ genes) showed a decrease in their expression toward the end of the gestation. The biological processes enriched in genes found in the maternal cluster that had an increasing expression pattern throughout the gestation were localization, transmembrane transport, regulation of smooth muscle cell migration, and positive regulation of cell morphogenesis (Dataset S12). The biological processes of the paternally expressed genes that had an increasing expression pattern throughout the gestation were macromolecule glycosylation, positive regulation of protein kinase B signaling, carbohydrate derivative metabolic process, and response to insulin. Maternal genes with expression that decreased throughout the gestation were predicted to be involved in proteolysis, the protein catabolic process, and the macromolecule catabolic process, while the paternal genes were involved in glucose transmembrane transport, telomere organization, telomere maintenance, regulation of DNA damage checkpoint, regulation of cell cycle checkpoint, and positive regulation of cell cycle.

Reciprocal Paternal and Maternal Gene Interaction. To further investigate the RNA–RNA interaction and functions of ASEGs, we compared them with the available Romiloski ligand-receptor pairs dataset that includes 2,557 ligand-receptor combinations (67). In total, 273 genes from the Romiloski dataset were detected in our dataset, including 19 genes we defined as being potentially imprinted (Dataset S13). We identified five putative interactions between the paternal and maternal genes and six ligand-receptor interactions which were transcribed from the same parents (Table 2). The expression patterns of these ligand-receptor genes were evaluated throughout gestation. Three reciprocal, parental ligand-receptor interactions showed a significant correlation in their expression pattern during pregnancy (Fig. 7). Maternally expressed *CALM2* showed an inverse correlation with the paternally expressed *INSR*, while *CALM2* and *IGF2R* (maternally expressed) showed a positive correlation with paternally expressed *ABCA1* and *IGF2*, respectively. Moreover, expression of four paternally expressed ligand-receptors (*IGF1:INSR*, *IGF2:INSR*, *IGF1:IGF1R*, and *DLK1:NOTCH4*) were positively correlated during pregnancy.

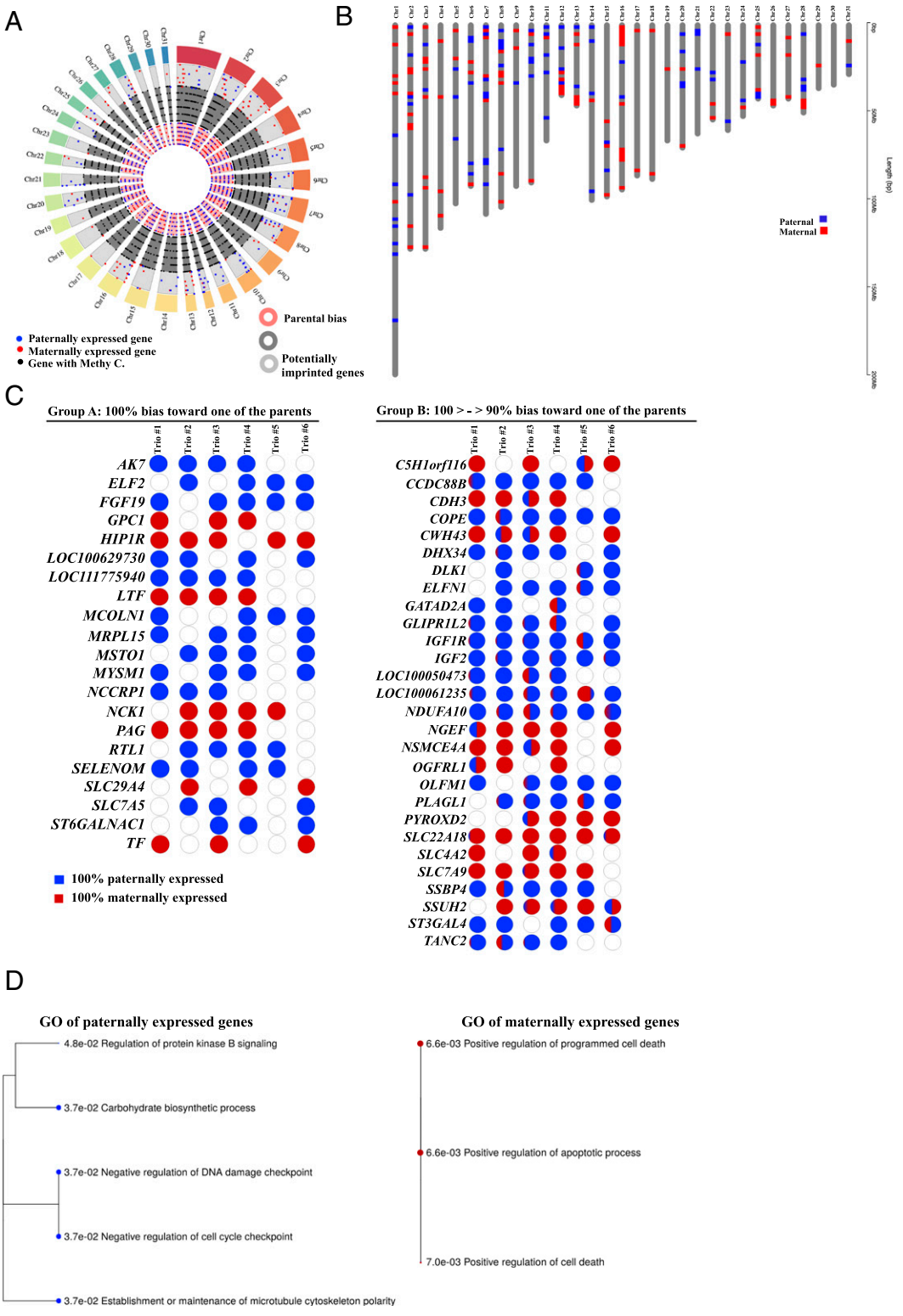

Fig. 4. (*A*) A circos plot of data curation to identify potentially imprinted genes. The list of genes with significant methylated cytosines, corresponding to 2,894 autosomal genes, was compared to the list of the ASEGs (n = 1,111). A total of 254 genes were both parentally expressed and have significant methylation site. Gene names can be found in the interactive circos plot at https://rpubs.com/pouyadini/568876. (*B*) An ideogram showing distribution of potentially imprinted genes on equine chromosomes. Gene names can be found in the interactive ideogram at https://rpubs.com/pouyadini/568783. The ideogram was generated using Package "ChromoMap" in R (62). (*C*) Presentation of parental bias in the expression of potentially imprinted genes (Groups A and B) within parent–placenta trios. (*D*) Hierarchical clustering tree of predicted GO of potentially imprinted genes. The size of the bullet points is associated with the number of genes involved in each process.

Since our RNA-seq reads were stranded, we were able to evaluate ASEGs, which are transcribed from an overlapping locus of another parentally expressed gene (opposite DNA strand). Overlapping loci were defined as with bidirectional gene expression within annotated genes (45). This definition allowed for the analysis of overlapping expression in allele-specific expressed loci. A total of 23 bidirectional ASEGs were identified, including six imprinted loci (Fig. 8*A*). We analyzed the correlation in the expression pattern of these overlapping loci throughout gestation (correlation between the expression of the sense and antisense strands). There were positive and negative correlations in the expression patterns of these overlapping loci (Table 3), with three overlapping loci showing transcription from opposite parent of origin (i.e., sense strand paternally and antisense strand maternally expressed). In the next step,

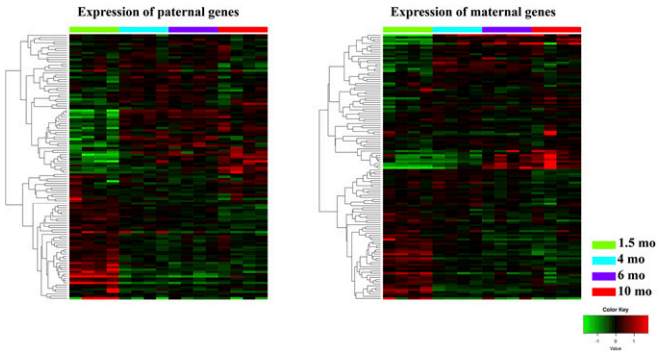

Fig. 5. Expression patterns of potentially imprinted genes during equine gestation. A total of 129 maternally (*Right*) and 125 paternally (*Left*) expressed genes were expressed during gestation. The heatmap was constructed using log(FPKM+1) values.

the position of each of these genes was visualized in the NCBI database for human and horse (human: GRCh38; horse: EquCab 3.0) to check the available data on the existence of overlapping loci. In the human and equine genome, there is evidence of overlapping expression for *LOC100051903* and *RTL1*. Based on the human and equine reference genomes, there were no signs of overlapping in *FAM184B*; however, we observed clear evidence of bidirectional expression in all these loci (Fig. 8*B*).

Discussion

ASEGs play an important role in placental and fetal development, making them interesting candidates for understanding maternal–fetal interactions and placental pathologies (47). Here, we described a workflow for analyzing ASE in the equine placental transcriptome and provided a list of genes with specific parent-of-origin expression. We identified 1,336 ASEGs (protein-coding and ncRNA), accounting for ∼4% of the current assembled genes in the equine genome. In the literature, the number of ASEGs may vary mainly based on defining cutoff points for maternal/paternal ratios. As mentioned in the results, different thresholds (from 60 to 100% bias toward one of the parents) have been used to determine the skew in the allelic expression (toward one of the parents). Also, there is currently no consensus over the number of SNPs per gene that is required to confirm significant paternal or maternal expression bias, with a percentage of monoallelic expressed (MAE) genes ranging from 2.4 to 10% between different studies in humans and mice using arbitrarily set thresholds (5, 69–72). In the current study, with a threshold of 75% expression bias toward one of the parents, the percentage of MAE genes was ∼4% (1,336 ASEGs out of 33,460 assembled genes in the equine genome), whereas ∼0.6% of the genes were identified as ASEGs when the threshold was set at 100% bias toward one of the parents (Group A, $n = 202$, including chrX; Dataset S3).

In the present dataset, we grouped the ASEGs based on the direction of the parental expression bias (Group A: 100%; Group B: 100 > to ≥ 90%; Group C: 90 > to ≥ 80%; and Group D: 80 > to ≥ 75%). We also determined the parental status of each gene within the individual trios. This approach allowed us to have a higher confidence for identifying a gene as an ASEG when it is expressed in all the trios with 100% skew to one of the parents (Group A). However, the other groups (B, C, and D) can also provide useful information when we take the number of SNPs in the gene (higher number of SNPs provides more confidence) and the parent of origin of the gene in each individual trio into consideration. Thus, information in this dataset can be interpreted with different levels of confidence and applied to diverse questions seeking to understand their dynamics in expression throughout placental development.

We also evaluated the distribution of ASEGs in the equine chromosomes. Surprisingly, ECA10 had a high number of

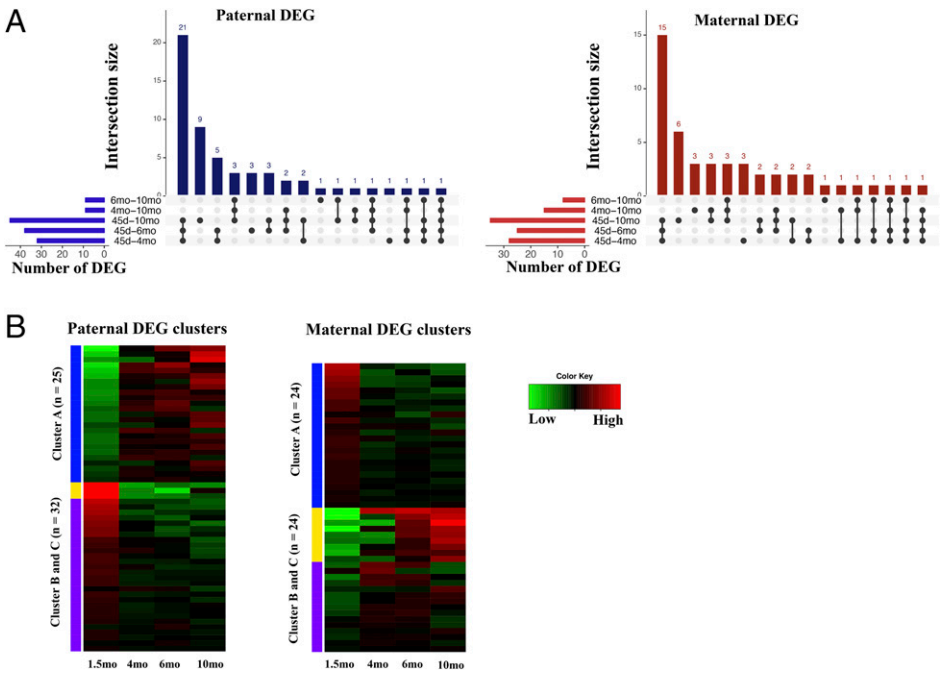

Fig. 6. (*A*) Differential expression of imprinted genes across gestation. Horizontal bars indicate the number of DEGs per comparison. Vertical bars indicate the number of DEGs shared across time points as indicated by black dots; connections indicate two or more conserved time points per transcript. No DEG was found between 4 and 6 mo. (*B*) Clustering of DEG based on their expression patterns during gestation. DEGs were clustered into three clusters using k-means clustering. The optimal number of clusters was evaluated with "mclust" package in R. The expression of two clusters consisting of 24 maternally expressed genes increased during pregnancy. There was also a cluster of maternally expressed genes in which the genes had higher expression at the beginning of the pregnancy and then declined toward the end of pregnancy. Within the paternally expressed genes, there was one cluster ($n = 25$ genes) which showed an increasing trend, and there were two clusters ($n = 32$ genes) that showed decreases in their expression toward the end of the gestation.

Table 2. Predicted ligand-receptor interactions among ASEGs

Ligand	Chromosome	Parental expression	Strand	Imprinted	Receptor	Chromosome	Parental expression	Strand	Imprinted	Correlation (FDR)
ADAM9	Chr27	Maternal	Forward	No	*ITGA6*	Chr18	Paternal	Forward	No	0.11 (>0.05)
CALM2	Chr15	Maternal	Reverse	Yes	*ABCA1*	Chr25	Paternal	Forward	Yes	0.62 (0.01)
CALM2	Chr15	Maternal	Reverse	Yes	*INSR*	Chr7	Paternal	Forward	Yes	−0.75 (<0.001)
DLK1	Chr24	Paternal	Reverse	Yes	*NOTCH4*	Chr20	Paternal	Forward	No	0.63 (0.009)
HGF	Chr4	Maternal	Forward	No	*ST14*	Chr7	Maternal	Reverse	Yes	0.28 (>0.05)
IGF2	Chr12	Paternal	Forward	Yes	*IGF2R*	Chr31	Maternal	Forward	Yes	0.92 (0.0001)
IGF1	Chr28	Paternal	Reverse	No	*IGF1R*	Chr1	Paternal	Forward	Yes	0.66 (0.006)
IGF1	Chr28	Paternal	Reverse	Yes	*INSR*	Chr7	Paternal	Forward	Yes	0.76 (0.0007)
IGF2	Chr12	Paternal	Forward	No	*IGF1R*	Chr1	Paternal	Forward	Yes	0.48 (>0.05)
IGF2	Chr12	Paternal	Forward	Yes	*INSR*	Chr7	Paternal	Forward	Yes	0.69 (0.0094)
SEMA6D	Chr1	Maternal	Reverse	No	*PLXNA1*	Chr16	Maternal	Reverse	Yes	0.06 (>0.05)

paternally expressed genes (52 paternally expressed [10 potentially imprinted] versus 8 maternally expressed [2 potentially imprinted]), and ECA3 had a high number of maternally expressed genes (63 maternally expressed [13 potentially imprinted] versus 14 paternally expressed [2 potentially imprinted]). ECA10 is the horse ortholog to human chromosomes 6 and 19, while ECA3 is the horse ortholog to human chromosomes 16 and 4 (73). Several imprinted coding and noncoding genes have been characterized on these human chromosomes (74–77). Interestingly, in humans, both coding and noncoding imprinted genes in these chromosomes have a substantial effect on the development and function of the placenta (74, 77, 78). In the present study, we propose a list of ASEGs and their chromosomal location; however, the function of these ASEGs in equine pregnancy needs to be further elucidated in future studies.

Imprinted genes are the best described class of ASEGs (5, 79). In addition, many of the known human imprinted genes are expressed in the placenta (80–82). The definite identification of imprinted genes in the equine placenta requires a more thorough assessment of the genome-wide methylation status of the parents and their placentae. Here, we performed RRBS on the CA samples and compared the list of ASEGs with the list of genes that have evidenced a significant increase in methylated cytosines. This approach led us to identify a set of potentially imprinted genes ($n = 254$) in the equine CA. Many of the already described imprinted genes were also identified in our dataset, including *IGF2*, *IGF2R*, *DLK1*, and *INSR*. The role of these genes in the development of the placenta has been extensively studied (9, 10, 83–85). It is worth considering that the absence of some of the known imprinted genes in our list could be due to the lack of expression of these genes in the equine placenta

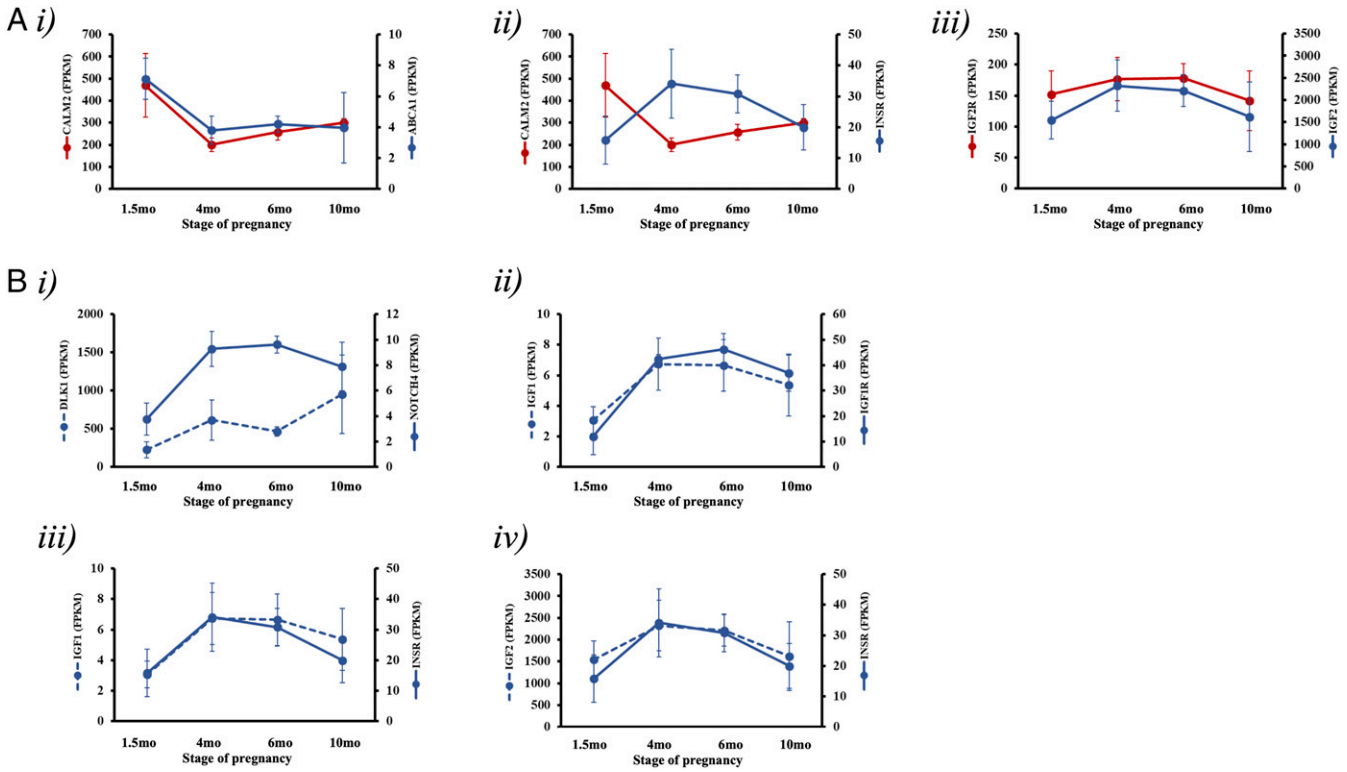

Fig. 7. The expression pattern of ligand-receptors throughout equine gestation. Seven ligand-receptor interactions showed significant correlation in their expression patterns throughout gestation. Three ligand-receptor interactions had a different parent of origin (*A*, *i*, *CALM2:ABCA1*; *ii*, *CALM2:INSR*; *iii*, *IGF2R:IGF2*) and four interactions were paternally expressed (*B*, *i*, *DLK1:NOTCH4*; *ii*, *IGF1:IGF1R*; *iii*, *IGF1:INSR*; *iv*, *IGF2:INSR*). The blue line indicates paternally expressed genes, and the red line indicates maternally expressed genes. The solid and dashed blue lines represent paternally expressed, ligand-receptor interactions.

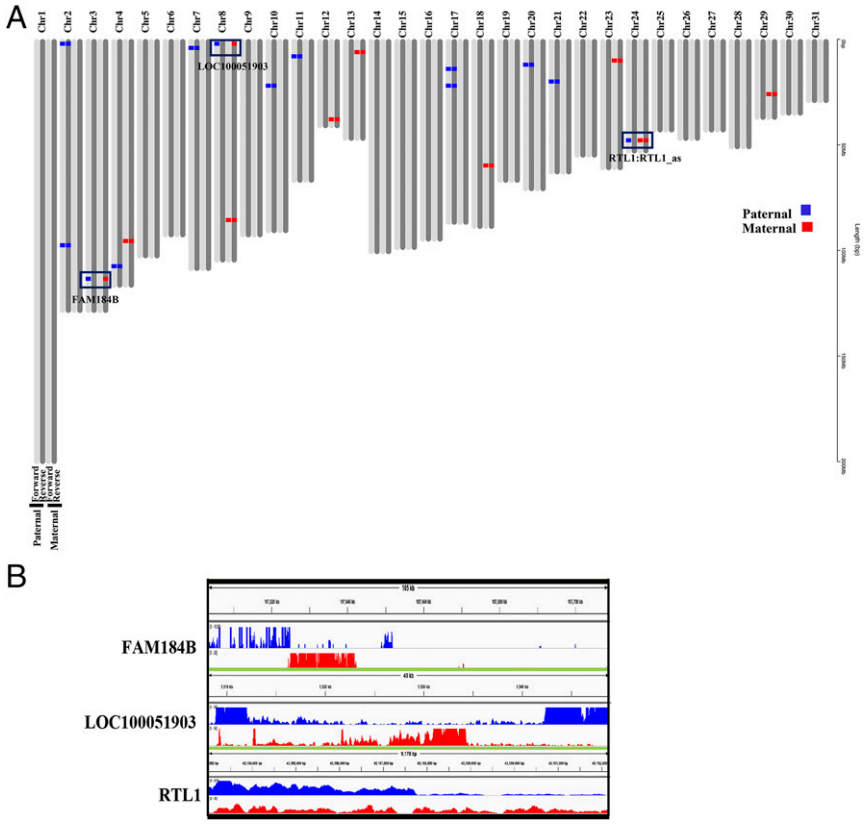

Fig. 8. (*A*) The location of overlapping loci throughout the equine genome. Each chromosome is indicated at the bottom. Overlapping loci were defined as bidirectionally expressed genes within annotated genes. A total of 50 ASE loci were identified with a bidirectional pattern of expression. Each dot represents a gene, and the horizontal bars indicate overlapping genes (red is maternally expressed, and blue is paternally expressed). The boxes indicate overlapping genes originated from different parents ($n = 3$). The ideogram was generated using Package "ChromoMap" in R (62). (*B*) The position of overlapping genes expressed from different parents of origin was visualized using the Integrative Genomics Viewer (68). Forward and reverse RNA-seq reads from trios were merged and converted to wig format. For each gene, the first plot represents the forward direction and the second plot (*Bottom*) represents the reverse direction. Blue indicates paternally expressed reads and red indicates maternally expressed.

(e.g., *INS*) or due to the absence of a sufficient number of informative SNPs in these genes (e.g., *H19*, *PHLDA2*, and *LY6G6C*). Exclusion of these genes reflects the strict criteria used to detect ASE in this study. However, we provided the data, and it can be reanalyzed with fewer restrictions. In addition, we used RRBS as a method for detecting methylated sites since it provides sufficient information on the methylome landscape of the samples. However, RRBS does not capture all CpG islands or promoters (86, 87). This could be seen on previously known imprinted genes, such as *PEG10* and *SNRPN*. These genes, among others, were identified as ASEGs, yet we were not able to confirm their methylation status in our dataset. It is important to mention that there was no discrepancy between the parent of origin in the previously known imprinted genes and ASEGs in our dataset. Additionally, the advantage of the current study is the separation of RNA-seq data based on their strand. Recently, a similar study showed a maternal bias in the expression of a well-known paternally expressed gene, *RTL1* (88). This discrepancy could be due to the presence of SNPs in the maternally expressed region, overlapping the *RTL1* gene (45). Since the RNA-seq data were not phased based on the strand, this sort of bias is expected. The information provided here adds to a growing body of knowledge in parental bias in placental gene expression.

In general, one of the accepted functions for ASE of genes in placenta is to provide a balance between the two opposing parental interests, leading to the development of a normal healthy pregnancy. In general, it is believed that there is a conflict between paternal and maternal genes over the level of maternal investment during pregnancy, a phenomenon known as parental conflict theory (16). The paternal genes expressed in the placenta promote the fetus to uptake more nutrients from the mother, while the maternal expressed genes attempt to save resources for the dam and maintain a balance in allocation of these resources (25–27). The interaction between ligands and receptors is one of the possible ways to achieve this balance. In the present study, we found three ligand-receptor combinations with different parent of origin, including *IGF2* and *IGF2R*. It is known that paternally expressed *IGF2* allocates maternal resources for the fetal benefit, and maternally expressed *IGF2R* binds and targets paternal *IGF2* for lysosomal degradation, thereby reducing its bioavailability (16, 27). Several studies have shown that alterations in the expression of *IGF2* and *IGF2R* are associated with placental and growth abnormalities (10, 89, 90). Analysis of the expression patterns of *IGF2* and *IGF2R* throughout gestation showed a positive correlation in the expression of these genes, suggesting an interaction between the paternal and maternal genes. There was also a significant correlation between maternally expressed *CALM2* and paternally expressed *ABCA1* and *INSR*. It is known that *INSR* binds to insulin, stimulates glucose uptake, and promotes fetal growth (91). However, there is no information about the role of *CALM2:INSR* and *CALM2:ABCA1* in the placenta, and the possible function of these interactions needs further investigation.

The other possible interaction between maternally and paternally expressed genes can be through sense and antisense

Table 3. Correlation analysis of the expression pattern between sense and antisense strands of overlapping ASEGs loci throughout gestation

Gene	Correlation (*r*)	FDR	Status	Imprinted
ALPK2	0.7319	0.0013	Overlapped	Yes
ATP7B	0.5548	0.0257	Overlapped	No
BRCA2	0.8287	0.0001	Overlapped	No
CELF2	0.3193	0.228	Overlapped	No
CPT1C	0.3517	0.1816	Overlapped	No
ELF2	0.3184	0.2294	Overlapped	Yes
EZH2	0.4954	0.051	Overlapped	No
FAM184B	−0.2891	0.1776	Overlapped from different parents of origin	No
LOC100051903	0.6288	0.0091	Overlapped from different parents of origin	No
LOC100063291	0.9925	0.0001	Overlapped	No
LOC100629324	0.543	0.0297	Overlapped	No
LOC100629730	0.4734	0.064	Overlapped	Yes
LOC106783107	0.9051	0.0001	Overlapped	No
MTREX	0.2491	0.3522	Overlapped	No
MYSM1	0.484	0.0575	Overlapped	Yes
NAP1L4	−0.3952	0.1298	Overlapped	No
NFE2L2	−0.8047	0.0131	Overlapped	No
PLPP1	0.9509	0.0001	Overlapped	No
PPFIA3	0.2245	0.4032	Overlapped	No
RASEF	0.8775	0.0001	Overlapped	No
RECQL5	−0.0634	0.8155	Overlapped	No
RTL1	−0.782	0.03998	Overlapped from different parents of origin	Yes
SLC29A4	0.9985	0.0001	Overlapped	Yes

transcription. It has been shown that antisense transcripts transcribed from the opposite strand of a particular gene locus can regulate their sense gene expression (41, 42). In this study, we found 23 overlapping loci through the equine genome which showed parental expression patterns. These sense–antisense transcriptions originate from a bidirectional expression of the genome (36–38, 40, 92–94). The majority of the overlapping loci in this study showed a positive correlation in their expression patterns. Most of these overlapping loci originated from one of the parents, except three of them, in which each strand originated from the opposite parent. One of the reciprocal paternal and maternal sense–antisense interactions in our dataset was *RTL1* and *RTL1_antisense*. It has been shown that the deletion of the paternal *RTL1* gene decreases the size of the murine placenta, while deletion of the maternal *RTL1_antisense* gene increases the size of the placenta (95, 96). It is proposed that these microRNAs located at the *RTL1_antisense* locus regulate the expression of *RTL1* in the mouse placenta (95, 97). We also found that impaired expression of *RTL1* is associated with insufficient function of the equine placenta (98). There is no information about the possible interaction of the other overlapping regions with reciprocal parental expression. The two other overlapping genes present in this study, which are expressed from opposite parent of origin, might be interesting candidates to evaluate their parental interaction for placental development during gestation.

The GO analysis of ASEGs and the potentially imprinted genes in this study further demonstrated potential interactions between maternally and paternally expressed genes in the placenta during the course of gestation. Our paternally expressed genes were predicted to be involved in metabolism and biosynthesis of protein and macromolecules, while maternally expressed genes were involved in localization and modification of these molecules. Moreover, maternally expressed genes were predicted to induce cell death and apoptosis in the placenta, a process that increases with placental growth and advancing gestation (99). This might suggest that the dam regulates the length of gestation by the expression of these genes. Similarly, several studies in humans also suggested that there is maternal control over the length of gestation; however, the exact mechanism has not been elucidated (100–102). In contrast to maternally expressed genes, we predicted that paternally expressed genes extend the length of gestation by delaying progression through the cell cycle and division until damaged DNA is repaired (103, 104). Overall, our data suggests that maternally expressed genes attempt to reduce the length of gestation, while paternally expressed genes attempt to extend it. Future studies to evaluate the contribution of ASEGs on the dynamic of placental development and its pathologies are warranted.

In conclusion, we have presented a dataset of ASEGs (publicly available at https://pouya-dini.github.io/equine-gene-db/) that includes 1,111 autosomal genes in the equine CA along with a workflow for analyzing allelic gene expression. We further identified 254 genes exhibiting ASE with significant methylated sites, thus being candidates of imprinted genes among these ASEGs. We predicted several reciprocal paternal and maternal interactions in gene expression in the equine CA. Lastly, we evaluated the expression pattern of the imprinted genes throughout gestation. Thus, this study provides fundamental information regarding parental gene expression during equine pregnancy, a species with a negligible amount of maternal contamination in its placenta. This information will aid further understanding of the role of parental bias in gene expression during gestation. Furthermore, these results may shed light on the gene regulatory processes underlying placental function, evolution, and disease. Future studies to understand the function of these genes and the reciprocal paternal and maternal interactions during placental development are warranted.

Materials and Methods

Detailed methodology is outlined in *SI Appendix, Supplementary Materials and Methods*.

Sample Collection, RNA and DNA Isolation, Whole Genome Sequencing, RNA-Seq, and RRBS of Parent–Offspring Trios. All animal procedures were prospectively approved by and completed in accordance with the Institutional Animal Care and Use Committee of the University of Kentucky (protocol numbers 2014-1341 and 2014-1215). Chorioallantoic samples were collected from pregnant mares at 4 mo ($n = 3$ [two females and one male]) and 6 mo ($n = 3$ [two males and one female]) of gestation. Paired whole blood was also collected form the sire and dam, respectively. RNA-seq (for CA), WGS (for CA and parental DNA), and RRBS (for CA) were performed on a HiSEq. 4000 (Illumina), generating six sets of parent–offspring trios.

Sequencing Data Processing and Bioinformatic Analysis. All the scripts for the analyses and their detailed descriptions can be found in the GitHub repository (https://github.com/P-Dini/PNAS-Dini2020). For the RNA-seq analysis, trimmed RNA reads were mapped to the reference horse genome (EquCab3.0) using STAR v2.5.2b (105) with a maximum of five mismatches allowed (–outFilterMismatchNmax 5). The mapped RNA reads were then phased based on the strand of transcription using the biwise FLAG in SAMtools v1.3.1 (F1R2 [Forward] and F2R1 [Reverse]) (106). For the WGS analysis, trimmed reads were aligned to reference genome (EquCab 3.0) using the Burrows–Wheeler algorithm (107). Variant calling on the mapped reads was performed using Genome Analysis ToolKit (GATK) in all samples (CA, dam, and sire) (106, 108). Generated variant call format (VCF) files of each trio were scanned to identify the genotype at each position in each trio and properly assign the parental origin of each variant found in the CA samples (script found in https://github.com/P-Dini/PNAS-Dini2020). SNPs at heterozygous sites which could clearly be assigned to one of the parents were considered informative and used for further analysis (*SI Appendix*, Fig. S1). These informative SNPs were matched to the corresponding gDNA sample (CA) to corroborate the correct assignment of each SNP (script found in https://github.com/P-Dini/PNAS-Dini2020). Finally, informative

SNPs were annotated using the reference equine transcriptome (EquCab3.0 GFF) by obtaining the overlapping genes for each variant according to the genomic coordinates (script found in https://github.com/P-Dini/PNAS-Dini2020). For the methylation analysis, RRBS raw reads were trimmed and mapped to the reference genome (EquCab3.0) using Bismark (109). The significant filtered methylated sites were annotated to the EquCab 3.0 reference transcriptome and grouped for each gene (script can be found in https://github.com/P-Dini/PNAS-Dini2020). DNA methylation was considered to be associated to a specific gene when found 1 kb upstream or downstream of the gene's transcription start site (TSS) or the transcription termination site (TTS), respectively (110, 111).

Identifying ASEGs. During the mapping of the RNA-seq and WGS reads to the reference genome (EquCab3.0), we used a Phred quality score of Q > 30 (base call with base call accuracy of 99.9%) and set six as the minimum number of SNPs per gene to be used for further analyses, which gives a probability of an incorrect call of 1×10^{-15}. In the variant calling process with GATK, we used a minimum confidence score threshold of Q30 with default parameters to define SNPs. Additionally, we removed SNPs with read depth below 10. A two-tailed binomial test was used to detect statistically significant deviations between the number of maternal and paternal SNPs for each gene under the assumption that parental alleles are expressed at equal levels (Pat = Mat = 0.5). Multiple testing-corrected *P* values were calculated by the false discovery rate (Benjamini and Hochberg—FDR < 0.05) method to control for false-positive results. To identify the significant ASEGs, we used a cutoff of maternal expression (Mat) > 75% and paternal expression (Pat) < 25% for maternally expressed candidates, and Mat < 25% and Pat > 75% for paternally expressed genes (8). Therefore, a gene required a minimum of 75% of the SNPs clearly assigned to a parent and an FDR < 0.05 to be considered as exhibiting allelic-specific expression. In addition to our analysis, we used the SNP-free RNA-editing identification toolkit (SPRINT) (112) to assess if our identified SNPs were located in regions prone to suffer RNA editing, using previously generated RNA-seq reads from CA (*n* = 16, GSE108279) (66) and the genome reference EquCab 3.0. Only 16 SNPs (0.004% of the total SNPs) were located in regions potentially affected by RNA editing. A total of 12 of these SNPs were located in 12 different genes (*LOC111772506*, *MAPKBP1*, *SCPEP1*, *KCNQ1*, *HDAC3*, *LOC100052533*, *ECE1*, *SCOC*, *SH3GLB2*, *COL18A1*, *FAM180A*, and *STAG2*), and 4 SNPs did not map to a gene. Importantly, the removal of these SNPs did not alter our subsequent analyses in the identification of ASEGs. Lastly, we investigated if our identified SNPs could exhibit a mappability bias, in which certain alleles have a better mapping rate than the alternative allele. For this purpose, we followed the WASP pipeline (113), which uses RNA-seq data (GSE108279) and removes reads unable to map to the exact same genomic location due to a change in the genotype of the SNP. In total, ~97% of the SNPs retained the same read depth or had five reads or less discarded. Further analysis using the same filters as previously described (including minimum read depth and number of SNPs per gene) determined no systematic mappability issue and no alteration to our subsequent analyses.

Data Availability. DNA sequencing and RRBS data have been deposited in Sequence Read Archive (PRJNA541840). The RNA-seq data from this study were deposited in the Gene Expression Omnibus (NCBI, NIH) database under the accession number GSE108279. All other study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We are grateful to Dr. James McLeod for his constructive comments. We are also thankful to Mr. Loring Bruck Daugherty at Gluck Genetic Laboratory for his assistance in DNA extraction. This work was funded by the Special Research Fund (Bijzonder Onderzoeksfonds; BOF) at Ghent University, the Albert G. Clay Endowment, and the Paul Mellon Postdoctoral fellowships at the University of Kentucky.

1. J. Rossant, J. C. Cross, Placental development: Lessons from mouse mutants. *Nat. Rev. Genet.* **2**, 538–548 (2001).
2. S. L. Adamson *et al.*, Interactions between trophoblast cells and the maternal and fetal circulation in the mouse placenta. *Dev. Biol.* **250**, 358–373 (2002).
3. R. Sood, J. L. Zehnder, M. L. Druzin, P. O. Brown, Gene expression patterns in human placenta. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5478–5483 (2006).
4. C. P. Gheorghe, R. Goyal, A. Mittal, L. D. Longo, Gene expression in the placenta: Maternal stress and epigenetic responses. *Int. J. Dev. Biol.* **54**, 507–523 (2010).
5. A. Gimelbrant, J. N. Hutchinson, B. R. Thompson, A. Chess, Widespread monoallelic expression on human autosomes. *Science* **318**, 1136–1140 (2007).
6. A. Chess, Monoallelic gene expression in mammals. *Annu. Rev. Genet.* **50**, 317–327 (2016).
7. A. C. Ferguson-Smith, Genomic imprinting: The emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12**, 565–575 (2011).
8. T. Metsalu *et al.*, Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics* **9**, 1397–1409 (2014).
9. R. Ono *et al.*, Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat. Genet.* **38**, 101–106 (2006).
10. C. P. Sibley *et al.*, Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8204–8208 (2004).
11. M. Charalambous *et al.*, Maternally-inherited Grb10 reduces placental size and efficiency. *Dev. Biol.* **337**, 1–8 (2010).
12. R. N. Plasschaert, M. S. Bartolomei, Genomic imprinting in development, growth, behavior and stem cells. *Development* **141**, 1805–1813 (2014).
13. W. R. Allen, Factors influencing pregnant mare serum gonadotrophin production. *Nature* **223**, 64–65 (1969).
14. W. P. Robinson, E. M. Price, The human placental methylome. *Cold Spring Harb. Perspect. Med.* **5**, a023044 (2015).
15. K. J. Jacob, W. P. Robinson, L. Lefebvre, Beckwith-wiedemann and Silver-russell syndromes: Opposite developmental imbalances in imprinted regulators of placental function and embryonic growth. *Clin. Genet.* **84**, 326–334 (2013).
16. J. M. Frost, G. E. Moore, The importance of imprinting in the human placenta. *PLoS Genet.* **6**, e1001015 (2010).
17. R. A. Fisher, M. D. Hodges, Genomic imprinting in gestational trophoblastic disease--a review. *Placenta* **24** (suppl. A), S111–S118 (2003).
18. S. Tomizawa, H. Sasaki, Genomic imprinting and its relevance to congenital disease, infertility, molar pregnancy and induced pluripotent stem cell. *J. Hum. Genet.* **57**, 84–91 (2012).
19. J. M. Kalish, C. Jiang, M. S. Bartolomei, Epigenetics and imprinting in human disease. *Int. J. Dev. Biol.* **58**, 291–298 (2014).
20. D. H. K. Lim, E. R. Maher, Genomic imprinting syndromes and cancer. *Adv. Genet.* **70**, 145–175 (2010).
21. Z. Chen, K. M. Robbins, K. D. Wells, R. M. Rivera, Large offspring syndrome: A bovine model for the human loss-of-imprinting overgrowth syndrome beckwith-wiedemann. *Epigenetics* **8**, 591–601 (2013).
22. P. Chavatte-Palmer *et al.*, Review: Placental perturbations induce the developmental abnormalities often observed in bovine somatic cell nuclear transfer. *Placenta* **33**, S99–S104 (2012).
23. B. Keverne, Monoallelic gene expression and mammalian evolution. *BioEssays* **31**, 1318–1326 (2009).
24. D. Monk *et al.*, Limited evolutionary conservation of imprinting in the human placenta. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6623–6628 (2006).
25. T. Moore, D. Haig, Genomic imprinting in mammalian development: A parental tug-of-war. *Trends Genet.* **7**, 45–49 (1991).
26. F. C. Cassidy, M. Charalambous, Genomic imprinting, growth and maternal-fetal interactions. *J. Exp. Biol.* **221**, jeb164517 (2018).
27. J. McMinn *et al.*, Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta* **27**, 540–549 (2006).
28. L. Korostowski, N. Sedlak, N. Engel, The Kcnq1ot1 long non-coding RNA affects chromatin conformation and expression of Kcnq1, but does not regulate its imprinting in the developing heart. *PLoS Genet.* **8**, e1002956 (2012).
29. D. Mancini-Dinardo, S. J. Steele, J. M. Levorse, R. S. Ingram, S. M. Tilghman, Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev.* **20**, 1268–1282 (2006).
30. F. Sleutels, R. Zwart, D. P. Barlow, The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810–813 (2002).
31. M. Girardot, J. Cavaillé, R. Feil, Small regulatory RNAs controlled by genomic imprinting and their contribution to human disease. *Epigenetics* **7**, 1341–1348 (2012).
32. J. Peters, J. E. Robson, Imprinted noncoding RNAs. *Mamm. Genome* **19**, 493–502 (2008).
33. J. E. Robson, S. A. Eaton, P. Underhill, D. Williams, J. Peters, MicroRNAs 296 and 298 are imprinted and part of the GNAS/Gnas cluster and miR-296 targets IKBKE and Tmed9. *RNA* **18**, 135–144 (2012).
34. H. Seitz *et al.*, A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res.* **14**, 1741–1748 (2004).
35. P. K. Yang, M. I. Kuroda, Noncoding RNAs and intranuclear positioning in monoallelic gene expression. *Cell* **128**, 777–786 (2007).
36. J. Shendure, G. M. Church, Computational discovery of sense-antisense transcription in the human and mouse genomes. *Genome Biol.* **3**, 10.1186/gb-2002-3-9-research0044 (2002).
37. B. Lehner, G. Williams, R. D. Campbell, C. M. Sanderson, Antisense transcripts in the human genome. *Trends Genet.* **18**, 63–65 (2002).
38. M. E. Fahey, T. F. Moore, D. G. Higgins, Overlapping antisense transcription in the human genome. *Comp. Funct. Genomics* **3**, 244–253 (2002).
39. G. G. Carmichael, *Antisense Starts Making More Sense* (Nature Publishing Group, 2003).
40. S. Katayama *et al.*; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium, Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564–1566 (2005).
41. M. Lapidot, Y. Pilpel, Genome-wide natural antisense transcription: Coupling its regulation to its different regulatory mechanisms. *EMBO Rep.* **7**, 1216–1222 (2006).
42. G. Lavorgna *et al.*, In search of antisense. *Trends Biochem. Sci.* **29**, 88–94 (2004).
43. S. Shibata, J. T. Lee, Tsix transcription- versus RNA-based mechanisms in Xist repression and epigenetic choice. *Curr. Biol.* **14**, 1747–1754 (2004).
44. K. Numata *et al.*, Identification of novel endogenous antisense transcripts by DNA microarray analysis targeting complementary strand of annotated genes. *BMC Genomics* **10**, 392 (2009).

45. P. Dini *et al*., Landscape of overlapping gene expression in the equine placenta. *Genes (Basel)* **10**, 503 (2019).
46. X. Wang, A. G. Clark, Using next-generation RNA sequencing to identify imprinted genes. *Heredity* **113**, 156–166 (2014).
47. S. Barbaux *et al*., A genome-wide approach reveals novel imprinted genes expressed in the human placenta. *Epigenetics* **7**, 1079–1090 (2012).
48. C. Daelemans *et al*., High-throughput analysis of candidate imprinted genes and allele-specific gene expression in the human term placenta. *BMC Genet.* **11**, 25 (2010).
49. H. Okae *et al*., Re-investigation and RNA sequencing-based identification of genes with placenta-specific imprinted expression. *Hum. Mol. Genet.* **21**, 548–558 (2012).
50. C. Proudhon, D. Bourc'his, Identification and resolution of artifacts in the interpretation of imprinted gene expression. *Brief. Funct. Genomics* **9**, 374–384 (2010).
51. P. L. Senger, *Pathways to Pregnancy and Parturition* (Current conceptions, INC, 2012).
52. X. Wang, D. C. Miller, R. Harman, D. F. Antczak, A. G. Clark, Paternally expressed genes predominate in the placenta. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10705–10710 (2013).
53. D. F. Antczak, W. R. Allen, Maternal immunological recognition of pregnancy in equids. *J. Reprod. Fertil. Suppl.* **37**, 69–78 (1989).
54. H. O. Hoppen, The equine placenta and equine chorionic gonadotrophin–An overview. *Exp. Clin. Endocrinol.* **102**, 235–243 (1994).
55. W. R. Allen, R. V. Short, Interspecific and extraspecific pregnancies in equids: Anything goes. *J. Hered.* **88**, 384–392 (1997).
56. T. S. Kalbfleisch *et al*., Improved reference genome for the domestic horse increases assembly contiguity and composition. *Commun. Biol.* **1**, 197 (2018).
57. C. M. Wade *et al*.; Broad Institute Genome Sequencing Platform; Broad Institute Whole Genome Assembly Team, Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **326**, 865–867 (2009).
58. A. Nag *et al*., Chromatin signature of widespread monoallelic expression. *eLife* **2**, e01256 (2013).
59. B. Reinius, R. Sandberg, Reply to 'High prevalence of clonal monoallelic expression'. *Nat. Genet.* **50**, 1199–1200 (2018).
60. B. Reinius *et al*., Analysis of allelic expression patterns in clonal somatic cells by single-cell RNA-seq. *Nat. Genet.* **48**, 1430–1435 (2016).
61. S. Vigneau, S. Vinogradova, V. Savova, A. Gimelbrant, High prevalence of clonal monoallelic expression. *Nat. Genet.* **50**, 1198–1199 (2018).
62. L. Anand, C. M. Rodriguez Lopez, chromoMap: An R package for interactive visualization and annotation of chromosomes. bioRxiv [Preprint] (2019). https://doi.org/10.1101/605600 (Accessed 3 January 2020).
63. S. Ge, D. Jung, R. Yao, G. O. Shiny, A graphical enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2019).
64. E. Habibi *et al*., Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).
65. C. A. Gifford *et al*., Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
66. P. Dini *et al*., Kinetics of the chromosome 14 microRNA cluster ortholog and its potential role during placental development in the pregnant mare. *BMC Genomics* **19**, 954 (2018).
67. J. A. Ramilowski *et al*., A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* **6**, 7866 (2015).
68. J. T. Robinson, H. Thorvaldsdóttir, A. M. Wenger, A. Zehir, J. P. Mesirov, Variant review with the integrative genomics viewer. *Cancer Res.* **77**, e31–e34 (2017).
69. C. Gregg *et al*., High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* **329**, 643–648 (2010).
70. L. M. Zwemer *et al*., Autosomal monoallelic expression in the mouse. *Genome Biol.* **13**, R10 (2012).
71. A.-V. Gendrel *et al*., Developmental dynamics and disease potential of random monoallelic gene expression. *Dev. Cell* **28**, 366–380 (2014).
72. S. M. Li *et al*., Transcriptome-wide survey of mouse CNS-derived cells reveals monoallelic expression within novel gene families. *PLoS One* **7**, e31751 (2012).
73. W. Li, J. Nadeau, E. Ostrander, B. Van Valkenburgh, P. Waddell, Comparative genomics: Mammalian radiations–genome maps 10. *Science* **286**, 463–478 (1999).
74. T. Eggermann *et al*., The maternal uniparental disomy of chromosome 6 (upd(6)mat) "phenotype": Result of placental trisomy 6 mosaicism? *Mol. Genet. Genomic Med.* **5**, 668–677 (2017).
75. J. Kim, L. Ashworth, E. Branscomb, L. Stubbs, The human homolog of a mouse-imprinted gene, Peg3, maps to a zinc finger gene-rich region of human chromosome 19q13.4. *Genome Res.* **7**, 532–540 (1997).
76. F. A. Middleton *et al*., Complete maternal uniparental isodisomy of chromosome 4 in a subject with major depressive disorder detected by high density SNP genotyping arrays. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **141B**, 28–32 (2006).
77. P. J. Yong, S. A. Marion, I. J. Barrett, D. K. Kalousek, W. P. Robinson, Evidence for imprinting on chromosome 16: The effect of uniparental disomy on the outcome of mosaic trisomy 16 pregnancies. *Am. J. Med. Genet.* **112**, 123–132 (2002).
78. D. M. Morales-Prieto, S. Ospina-Prieto, W. Chaiwangyen, M. Schoenleben, U. R. Markert, Pregnancy-associated miRNA-clusters. *J. Reprod. Immunol.* **97**, 51–61 (2013).
79. M. F. Lyon, X chromosomes and dosage compensation. *Nature* **320**, 313 (1986).
80. F. Court *et al*., Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res.* **24**, 554–569 (2014).
81. M. Sanchez-Delgado *et al*., Absence of maternal methylation in biparental hydatidiform moles from women with NLRP7 maternal-effect mutations reveals widespread placenta-specific imprinting. *PLoS Genet.* **11**, e1005644 (2015).
82. C. W. Hanna *et al*., Pervasive polymorphic imprinted methylation in the human placenta. *Genome Res.* **26**, 756–767 (2016).
83. A. Smallwood *et al*., Temporal regulation of the expression of syncytin (HERV-W), maternally imprinted PEG10, and SGCE in human placenta. *Biol. Reprod.* **69**, 286–293 (2003).
84. A. Yevtodiyenko, J. V. Schmidt, Dlk1 expression marks developing endothelium and sites of branching morphogenesis in the mouse embryo and placenta. *Dev. Dyn.* **235**, 1115–1123 (2006).
85. J. St-Pierre *et al*., IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics* **7**, 1125–1132 (2012).
86. H. Guo *et al*., Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat. Protoc.* **10**, 645–659 (2015).
87. A. Meissner *et al*., Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
88. D. Pilvar, M. Reiman, A. Pilvar, M. Laan, Parent-of-origin-specific allelic expression in the human placenta is limited to established imprinted loci and it is stably maintained across pregnancy. *Clin. Epigenetics* **11**, 94 (2019).
89. A. Gabory, M.-A. Ripoche, T. Yoshimizu, L. Dandolo, The H19 gene: Regulation and function of a non-coding RNA. *Cytogenet. Genome Res.* **113**, 188–193 (2006).
90. M. M. Lau *et al*., Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes Dev.* **8**, 2953–2963 (1994).
91. A. Efstratiadis, Genetics of mouse growth. *Int. J. Dev. Biol.* **42**, 955–976 (1998).
92. H. Kiyosawa, N. Mise, S. Iwase, Y. Hayashizaki, K. Abe, Disclosing hidden transcripts: Mouse natural sense-antisense transcripts tend to be poly(A) negative and nuclear localized. *Genome Res.* **15**, 463–474 (2005).
93. J. Chen, M. Sun, L. D. Hurst, G. G. Carmichael, J. D. Rowley, Genome-wide analysis of coordinate expression and evolution of human cis-encoded sense-antisense transcripts. *Trends Genet.* **21**, 326–329 (2005).
94. S. H. Munroe, J. Zhu, Overlapping transcripts, double-stranded RNA and antisense regulation: A genomic perspective. *Cell. Mol. Life Sci.* **63**, 2102–2118 (2006).
95. M. Ito *et al*., A trans-homologue interaction between reciprocally imprinted miR-127 and Rtl1 regulates placenta development. *Development* **142**, 2425–2430 (2015).
96. M. Kitazawa, M. Tamura, T. Kaneko-Ishino, F. Ishino, Severe damage to the placental fetal capillary network causes mid- to late fetal lethality and reduction in placental size in Peg11/Rtl1 KO mice. *Genes Cells* **22**, 174–188 (2017).
97. E. Davis *et al*., RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr. Biol.* **15**, 743–749 (2005).
98. P. Dini *et al*., Equine hydrallantois is associated with impaired angiogenesis in the placenta. *Placenta* **93**, 101–112 (2020).
99. A. N. Sharp, A. E. Heazell, I. P. Crocker, G. Mor, Placental apoptosis in health and disease. *Am. J. Reprod. Immunol.* **64**, 159–169 (2010).
100. M. Jainudeen, E. Hafez, "Gestation, prenatal physiology, and parturition" in *Reproduction in Farm Animals*, B. Hafez, E. S. E. Hafez, Eds. (Wiley-Blackwell, ed. 7, 2000), pp. 140–155.
101. G. Zhang *et al*., Genetic associations with gestational duration and spontaneous preterm birth. *N. Engl. J. Med.* **377**, 1156–1167 (2017).
102. J. B. Wu, J. Sha, J. F. Zhai, Y. Liu, B. Zhang, Prenatal diagnosis of maternal partial trisomy 9p23p24.3 and 14q11.2q21.3 in a fetus: A case report. *Mol. Cytogenet.* **13**, 6 (2020).
103. L. Wang, H. Chen, C. Wang, Z. Hu, S. Yan, Negative regulator of E2F transcription factors links cell cycle checkpoint and DNA damage repair. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3837–E3845 (2018).
104. I. A. Shaltiel, L. Krenning, W. Bruinsma, R. H. Medema, The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle. *J. Cell Sci.* **128**, 607–620 (2015).
105. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
106. H. Li *et al*.; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
107. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
108. A. McKenna *et al*., The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
109. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
110. H. Gu *et al*., Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468–481 (2011).
111. J. A. Gim *et al*., Genome-wide analysis of DNA methylation before-and after exercise in the thoroughbred horse with MeDIP-Seq. *Mol. Cells* **38**, 210–220 (2015).
112. F. Zhang, Y. Lu, S. Yan, Q. Xing, W. Tian, SPRINT: An SNP-free toolkit for identifying RNA editing sites. *Bioinformatics* **33**, 3538–3548 (2017).
113. B. van de Geijn, G. McVicker, Y. Gilad, J. K. Pritchard, WASP: Allele-specific software for robust molecular quantitative trait locus discovery. *Nat. Methods* **12**, 1061–1063 (2015).
114. C. Trapnell *et al*., Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
115. C. Trapnell *et al*., Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
116. P. Ruzzenenti *et al*., The Ferritin-Heavy-Polypeptide-Like-17 (FTHL17) gene encodes a ferritin with low stability and no ferroxidase activity and with a partial nuclear localization. *Biochim. Biophys. Acta* **1850**, 1267–1273 (2015).
117. S. Kobayashi *et al*., The X-linked imprinted gene family Fthl17 shows predominantly female expression following the two-cell stage in mouse embryos. *Nucleic Acids Res.* **38**, 3672–3681 (2010).