An investigation of mechanisms underlying mouse blastocyst hatching: an RNA sequencing study

Running title: The mechanism of hatching

Yaqiong Liu, PhD; Celine Jones; Kevin Coward, PhD*

Nuffield Department of Women's and Reproductive Health, University of Oxford, Level 3, Women's Centre, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK.

*Corresponding author

Email: kevin.coward@wrh.ox.ac.uk; Telephone: +44 [0] 1865 618900; Fax: +44 [0] 1865 618900

Declarations of interest: none

ABSTRACT

Objective: To investigate the regulatory mechanisms and signaling molecules underlying hatching in mouse embryos. ~~hatching.~~

Design: Experimental laboratory study using a mouse embryo model.

Setting: University-based basic scientific~~ce~~ research laboratory.

Animals: A total of 40 B6C3F1 × B6D2F1 mouse embryos were used in this study.

Interventions: Frozen/thawed mouse embryos, at the 8-cell stage, were cultured *in vitro* for two days. The resulting hatching and pre-hatching blastocysts were then used for cDNA library preparation and RNA-sequencing analysis (n=8 for each group). ~~The~~ Didifferentially expressed genes (DEGs) were ~~further used for~~then used for downstream functional analysis. In addition, a list of genes related to developmental progression in humans was used to identify genes that were potentially related to the hatching of human embryos.

Main Outcome Measures: ~~Differentially expressed genes (DEGs),~~DEGS, enriched gene ontology (GO) terms and canonical pathways, clustered gene~~s~~ networks, activated upstream regulators, and ~~the~~ common genes between a gene list of ~~mouse~~ hatching-related genes in mice ~~list~~ and a gene list associated with ~~human~~ developmental progression ~~gene list.~~in humans.

Results: A total ~~of 275 differentially expressed genes (DEGs)~~ DEGs were identified between hatching and pre-hatching blastocysts: 230 up-regulated genes and 45 down-regulated genes (adjusted $P<0.05$). Functional enrichment analysis suggested

that blastocyst hatching *in vitro* is an ATP-dependent process that involves protein biosynthesis and organization of the cytoskeleton. Furthermore, by regulating cell motility, the RhoA signaling pathway (including *Arpc2, Cfl1, Gsn, Pfn1, Tpi1, Grb2, Tmsb10, Enah, and Rnd3* genes) may be a crucial signaling pathway during hatching. Furthermore, we identified a cluster of genes (*Krt8, Krt7, Cldn4*, and *Aqp3*) that exerted functional roles in cell-cell junctions and water homeostasis during hatching. Moreover, some growth factors (AGT and FGF2) and endocrine factors (oestrogen receptor and prolactin), were predicted to be involved in the regulation of embryo hatching. In addition, we identified 81 potential genes that are potentially involved in the hatching process in human embryos.

Conclusion: Our analysis identified potential genes and molecular regulatory pathways involved in the blastocyst hatching process in mice; we also identified genes that may potentially regulate hatching in human embryos. Our findings enhance our knowledge of embryo development and provides useful information for further exploring the mechanisms underlying embryo hatching.

Keywords: Mouse embryo; hatching; RNA sequencing; mechanism; pathway; upstream regulator

Introduction

Blastocyst hatching is a prerequisite for embryo implantation and is accompanied by partial rupture of the zona pellucida (ZP) and subsequent dynamic movements of the embryo into the external environment (1). The structure of the ZP is comparable across most species (2), however, the hatching process is a highly intriguing biological phenomenon. Previous studies have suggested that hatching is the result of a well-orchestrated sequence of increasing mechanical pressure, protease activation, and the dynamic formation of trophectodermal projections (TEPs). Furthermore, it is hypothesised that hormones, and the secretion of growth factors, also play a key role and that this physiological process may be controlled *via* cross-talk between embryos and the maternal uterus (3-8).

In most mammals, blastocyst hatching is thought to be subjected to both mechanical and chemical factors (5). First of all, the ZP can be affected by the hydrostatic pressure exerted by the expanded blastocyst; the ZP is subsequently stretched and becomes thinner (6). It has been demonstrated that the formation of the blastocyst cavity is closely regulated by active ion transport (Na^+/K^+-ATPases), aquaporin water channels (AQPs), and tight junctions (such as tight junction proteins (TJPs), occludins, and claudins) (9-13). In addition, a previous study found that the inhibition of Na^+/K^+-ATPase reduced the probability of hatching (6). The roles of these molecules have been well established during the formation of the blastocyst cavity; however, it remains unclear as to whether they are directly involved in the hatching process.

Apart from mechanical force, the enzymatic degradation of the ZP is believed to play an important in the hatching process; however, different classes of protease, such as serine proteases, cysteine proteases (cathepsin), and matrix metalloproteinases

(MMP), are known to participate in the hatching in different species (1,14,15). Furthermore, a significant number of growth factors (such as HB-EGF, LIF, and EGF), hormones (e.g., progesterone), and cytokines (e,g., IL-6, GM-GSF, and IL-11), have all been found to play a role in hatching *via* the crosstalk between embryos and the maternal uterus (1,16-19). However, most previous studies targeted only a limited number of molecules in each experiment by using related inhibitors or activators. Nevertheless, the regulatory mechanisms of these maternally derived or embryo-derived secretions on embryo hatching remains unknown.

Trophectodermal projections (TEPs) are single, finger-like, F-actin enriched cytoplasmic outgrowths. The formation of these TEPs has been observed during the hatching process in many species (1,5,20). It is widely believed that TEPs can participate in the initiation of the hatching process as well as embryo attachment (5). Furthermore, some studies have indicated that cytoskeleton organization, mostly relating to actin-based modifications (F-actin formation) in TE cells, potentially take parts in shaping the blastocysts to hatch from the ZP (5,7). However, studies related to the biogenesis of TEPs, embryo cell movement, and cytoskeleton organization, are sparse. It is not yet clear which active molecules and related biological signalling pathways underly this process.

Collectively, it is highly likely that multiple molecules and regulatory mechanisms are involved in the blastocyst hatching process; however, to the best of our knowledge, the patterns of embryonic gene expression during this short period of embryo development remain elusive. Given the importance of this physiological process in embryo development and the establishment of pregnancy, it is necessary to profile the transcriptome of hatching embryos, and to examine hitherto unstudied signaling

mechanisms that might regulate embryo hatching. Furthermore, although clinical observations have indicated that human embryos begin hatching between Day 5 and Day 7 after fertilization; most investigations regarding embryo hatching mechanisms have been carried out in animal models. Therefore, it is imperative that we focus on hatching mechanisms in human embryos.

Advancements in high-throughput next-generation sequencing (NGS) techniques have created powerful tools for studying physiological and pathophysiological phenomena. Of these NGS techniques, RNA-sequencing (RNA-seq) has become the standard technology for the transcriptomic analysis of genomics and is often used to analyze differential gene expression due to its high degree of accuracy, extension, and affordability (21,22). Previously, researchers have used RNA-seq to improve our understanding of the embryonic transcriptome, particularly with regards to the molecular mechanisms involved in embryo linage and differentiation (23-26). More recently, with the development of the single-cell RNA-seq technique, it is possible to investigate the role of different cell populations during embryogenesis (27-30). However, information regarding the molecular regulation of hatching remains largely elusive.

In order to address the clear gaps in knowledge relating to this critical stage of embryo development, we used RNA-seq and downstream analysis to characterize the transcriptional profile of hatching blastocysts in a mouse model to explore the potential mechanistic pathways underlying the hatching process. In addition, to explore the potential genes and pathways involved in the hatching of human embryos, we compared our dataset with a published dataset of genes related to developmental

progression in human embryos and identified common differentially expressed genes (DEGs).

Methods and materials

Embryo preparation and culturing

Cryopreserved 8-cell stage mouse embryos (B6C3F1 × B6D2F1) were purchased from Embryotech Laboratories (Wilmington, MA, USA). Each cryopreserved straw contained 20 embryos which were randomly collected from embryo-pools harvested from 8 to 12 mice. The batch number of the straw was marked and identified according to the origin of the contained embryos. Generally, four straws of frozen embryos were thawed in accordance with the instructions provided by the supplier (4 x 20 = 80 embryos); half of these embryos were used for experimental purposes (n=40). After thawing, the embryos were transferred into KSOM-AA medium (Millipore, Abingdon, UK). Embryos were cultured individually in KSOM-AA medium and covered with mineral oil (FertiPro NV, Beernen, Belgium) at 37°C with 5% CO_2. According to a previous study, the mean time of hatching commencement was 108 hours post-hCG injection (E4.5) for *in vitro* developed embryos (31); therefore, we cultured embryos *in vitro* for 48 hours to E4.5; at this point, some of the embryos would hatch spontaneously while the others would be at the expansion stage.

Library preparation and RNA sequencing

After 48 hours of *in vitro* culture, embryos were assessed morphologically. Analysis showed that 90.0% (36/40) of the embryos developed to the blastocyst stage and only 36.1% (13/36) of embryos reached the hatching stage. Representative images of the blastocysts in the hatching group and pre-hatching group are shown in Supplementary

Fig 1. Overall, 8 hatching blastocysts and 8 pre-hatching expanded blastocysts were randomly collected for subsequent RNA sequencing experiments. Whole single blastocysts were lysed in 2.3 µl of lysis buffer (0.8% (vol/vol), Triton X-100, and 2 U/µl of RNase inhibitor (both from Thermo Fisher, UK)) for cDNA synthesis using the SmartSeq2 method, as described previously (32). Libraries were then prepared using using Nextera-XT (Illumine, US). ERCC RNA Spike-in mix (ThermoFisher, UK) was added at a 1/100,000 dilution. Amplified libraries were analysed for size distribution using the TapeStation High Sensitivity D1000 kit (Agilent, US). Libraries were then quantified using PicoGreen (Invitrogen, US) and relative volumes were pooled accordingly. Sequencing was then performed as 75bp paired-end reads on a HiSeq4000 system according to Illumina specifications.

RNA-sequencing (RNA-seq) data processing and quality control

RNA-seq reads were aligned to the mouse reference genome (GRCm38) using HISAT2 (33) and duplicate reads were removed using the Picard 'MarkDuplicates' tool (http://broadinstitute.github.io/picard). Reads mapping uniquely to Ensembl-annotated genes (approximately 5 million per sample) were summarised using FeatureCounts (34). The raw gene count matrix was imported into the R/BioConductor environment (35) for further processing and analysis. Advanced RNA quality control was then conducted to test the similarity and distance between library samples: (1) read assignment by category, (2) count metrics, (3) gene detection curve, (4) read biotype counts, and (5) sample clustering. Two sample libraries in the pre-hatching group resulted in a low gene detection level (Supplementary **Fig 2A**). Furthermore, Spearman's correlation test showed these two samples were positioned further away from other samples (Supplementary **Fig 2C**); therefore, these samples were

considered unsatisfactory and were excluded from further processing (Supplementary **Fig 2 B** and **D**). Further steps included data normalisation and applying a filter to include only genes with at least 10 reads in at least 3 of the samples.

Differentially expressed genes (DEGs) analysis

The *edgeR* package (36) was used to test for differentially expressed genes between hatching blastocysts and pre-hatching blastocysts. This package combines the use of the trimmed mean of M-values as the normalisation method for count data, the empirical Bayes approach for estimating tagwise negative binomial dispersion values, and finally, generalized liner models and quasi-likelihood F-tests for detecting DEGs for the hatching stage. Embryo batch was also included as an additional explanatory variable. Raw *P*-values were adjusted for multiple testing using the Benjamini-Hochberg (BH) procedure (37) to control the false discovery rate at 5%.

Functional enrichment analysis

All DEGs were analysed using Over Representation Analysis (ORA) to identify enriched Gene Ontology (GO) terms (38). These enrichment analyses were performed in the R environment using the *ClusterProfiler* (39) package. In addition, the canonical pathway analysis and upstream regulator analysis were generated through the use of Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis).

Construction and analysis of the protein-protein network

Protein-protein interactions (PPI) networks that corresponded to the DEGs were retrieved from the Search Tool for Retrieval of Interacting Genes (STRING) database (40). The confidence score cut-off value was set as 0.4, and the PPI network was

visualized by Cystoscope software (41). Furthermore, Molecular Complex Detection (MCODE) (42) was used to investigate the substructure of the PPI network extracted from the constructed network and focused on highly connected nodes (known as clusters).

Validation

To corroborate gene expression from the RNA-seq analysis, twelve DEGs and one unchanged gene were selected for validation by quantitative real-time PCR (qRT-PCR). Embryo cDNA was directly synthesised from three whole blastocysts using a cell-to-cDNA kit (Invitrogen, UK). Because of the small amount of cDNA produced from the three embryos, we pre-amplified the target cDNA using the TaqMan PreAmp Mater Mix Kit (Applied Biosystems, UK), in accordance with the manufacturer's instructions. Pooled TaqMan Assays were then created by combining an equal volume of each 20 × TaqMan probe: (*Actb* (Mm02619580_g1), *Rnd3* (Mm00512162_m1), *Myl6* (Mm02342525_g1), *Cox8a* (Mm02342396_g1), *Tmeff1* (Mm05907935_s1), *Ctsh* (Mm00514455_m1), *Lgals1* (Mm00839408_g1), *Krt18* (Mm01601704_g1), *Aqp3* (Mm01208559_m1), *Atp5b* (Mm01160389_g1), *Rpl13a* (Mm05910660_g1), *Ndufa7* (Mm00458227_m1), and *Cdx2* (Mm01212280_m1)). RT-qPCR was then carried out using the TaqMan Gene Expression System (Applied Biosystem, UK), running an Applied Biosystems QuantStudio 3 (Thermo Fisher, UK) with standard thermocycler parameters. After validating appropriate housekeeping genes, the mean CT (cycle threshold) values for *Canx* and *Ywhaz* were used as the internal controls. Subsequently, the relative expression levels of target genes were calculated using the $2^{(-\Delta\Delta Ct)}$ method (43). The expression of target genes was firstly normalised to the internal control; then, the relative gene expression in the treatment/ hatching group

was presented as the fold change relative to the mean Ct value of the control/pre-hatching group ($2^{-(\text{CT treatment-CT control})}$). The validation experiment was conducted with six biological repeats. Relative gene expression levels were compared between hatching and pre-hatching groups using the student's t-test. In addition, correlation analysis was used to investigate the association of fold changes in gene expression were compared between datasets obtained by RNA-seq and RT-PCR. A probability (*P*) value < 0.05 was considered to be statistically significant.

Common gene identification between mouse hatching and human developmental progression

A dataset of genes related to developmental progression in human embryos was previously generated by Petropoulos et al. (30) by single cell RNA-seq. In this dataset, DEGs were identified by comparing between different embryonic days and between different time-points in the same linage. A significance cut-off of 0.05 (adjusted) was used to determine whether a gene was significantly expressed at a developmental timepoint. Considering that human hatching is initiated on Day 5 to Day 7, we grouped the DEGs between E5 and E6, and between E6 and E7. The identification of common genes between the dataset of genes related to human development and genes related to mouse hatching were obtained using the R programming language. Downstream analysis (described above) was also conducted to identify terms and pathways that were enriched.

Results

According to the advanced quality control results, we used eight hatching blastocysts and six pre-hatching blastocysts in our subsequent analyses. After normalisation of

the data, and filtration of the genes with low expression levels, a total of 10861 assigned genes passed the filter criteria and were used for DEG analysis.

Differentially expressed genes in hatching E4.5 blastocysts

Alternations in the blastocyst transcriptome at the hatching stage were analysed by comparing hatching *vs* pre-hatching blastocysts, and 275 DEGs were identified (BH adjusted P<0.05); of these, 230 were up-regulated and 45 were down-regulated (Supplementary **Table 1**). However, the absolute fold changes of these DEGs were relatively small; most were smaller than 2 (logFC <1) (**Fig.1A)**. The expression profiles of these DEGs were then extracted and displayed as a clustered heatmap; the sample clustering is shown in **Fig.1B** with red representing pre-hatching samples and green representing hatching samples at the top of the plot**.** In addition, we compared these DEGs with previously published mouse linage-specific sing-cell RNA seq datasets. We found that there were 16 trophectoderm (TE)-specific genes (including *Lgals1, Krt18, Actb, Myl6, Ctsh, and Krt8.*) and six genes that were specific to the inner cell mass (*Rnd3, Tmeff1, Rpl13a, Qki, Rps20, Myef2*). Interestingly, all these TE-specific genes were up-regulated in hatching blastocysts (**Fig.1C**).

The downstream functional analysis of DEGs

The DEGs in hatching blastocysts were significantly enriched in 33 GO terms for biology process, 18 GO terms for molecular function, and 14 GO terms for cell components (Supplementary **Table 2**). As shown in **Fig. 2A and** Supplementary **Fig 3**, these enriched biological terms were related to three main process and functions: (1) ribosome biogenesis; (2) ATP (adenosine triphosphate) metabolic process, including energy coupled proton transport and down electrochemical gradient; and (3) regulation of cytoskeleton organization and actin polymerization. Furthermore, apart

from the ribosome and cytochrome, other cell components, such as keratin filament and podosome were also enriched (Supplementary **Fig 3**).

Furthermore, there were 63 canonical pathways7 enriched KEGG pathways, 3 Wiki pathways, 17 Reactome pathways showing enrichment among these DEGs (adjusted *P*<0.05) (Supplementary **Table 3**). The top-ranked enriched pathways (-log (p-value)>2) are listed in **Fig 3B.** These pathways were mainly related with translation-related signaling (EIF2 signaling and regulation of eIF4 and p70S6K signalling), metabolic-related signaling (oxidative phosphorylation and the sirtuin signalling pathway), and actin cytoskeleton signalling (mTOR signalling, regulation of actin-based motility by Rho, RhoA signalling, and actin cytoskeleton signalling). In addition, canonical pathway analysis demonstrated that EIF2 signaling (Z-score=4.47), oxidative phosphorylation (Z-score=3.74), the regulation of actin-based motility by Rho (Z-score= 1.34), and the RhoA signalling pathway (Z-score=2.64) (related molecules included *Actb, Arpc2, Cfl1, Myl6, Pfn1, Rnd3, and Septin9)* were predicted as being activated in hatching blastocysts. In contrast, the RhoGDI (Rho GDP-dissociation inhibitor) signaling pathway (Z-score=-1) was inhibited (**Fig.2B**).

The predicted upstream regulator involved in hatching

The goal of IPA upstream regulator analysis was to identify the cascade of upstream transcriptional regulators that can explain the observed changes in gene expression and can help to identify the biological activities occurring in blastocysts. According to IPA upstream regulator analysis, 23 upstream regulators (with an activation Z-score>2) were shown to be activated in the hatching blastocysts. Transcription regulators, such as MYC (MYC proto-oncogene, BHLH transcription factor), TP53 (tumor protein P53), RB1 (RB transcriptional corepressor 1), and Esrra (estrogen-related receptor alpha),

were predicted to be activated in the hatching blastocysts. In addition, growth factors, including (AGT (angiotensinogen), and FGF2 (fibroblast growth factor 2) were also predicted to be activated. Furthermore, oestrogen receptor (*Cldn4, Krt18, Krt7, Krt8,* and *Tmeff1*), prolactin (PRL) (*Actb, Atp5pd, Cox4i1, Ctsh, Tmsb4x,* and *Tuba1b*), insulin, and IGF1R (insulin like growth factor 1 receptor), were all found to be activated according to their the status of their regulatory genes. Furthermore, nine upstream regulators (with an activation Z-score<-2) were shown to be inhibited in hatching blastocysts (**Fig.2C**). The upstream regulators and target molecules are presented in Supplementary **Table 4.**

The PPI network and modules

Based on the String database, a PPI network featuring 167 nodes and 1866 edges was generated from the 275 DEGs. To identify potential interactions between these molecules, we used MCODE to identify densely connected networks; four modules were shown to be significantly clustered (**Fig.3**). Module A included 103 nodes and 1805 edges (MCODE score: 35.4), while module B included 35 nodes and 176 edges (MCODE score: 10.4). The molecules in these modules were enriched in the pathways and terms related to ribosome biogenesis, oxidative phosphorylation, and the actin cytoskeleton. In addition, module C (MCODE score: 2.9) included 12 nodes and 16 edges (*Cfl1, Arpc2, Gsn, Pfn1, Shmt1, Tuba1b, Tmsb10, Tpi1, Ldlrap1, Rnd3, Enab, and Grb2*); these molecules were related to the regulation of the actin cytoskeleton pathway (*Cfl1, Arpc2, Gsn, Pfn1, and Enab*), the RHO GTPase effectors pathway (*Tuba1b, Pfn1, and Arpc2*); other GO terms were related to actin filament organisation. Furthermore, module D included *Krt8, Krt7, Krt18, Cldn4*, and *Aqp3* (MCODE score: 3.0). These molecules were related with filaments (*Krt8, Krt7,* and *Krt18*), cell-cell

junctions (*Krt8, Cldn4,* and *Aqp3*), channel activity, and water homeostasis (*Cldn4* and *Aqp3*).

Validation of RNA-seq result

To validate RNA-seq results, we selected certain genes based on their significance and the implicated pathways, including genes related to the actin-based cytoskeleton (*Actb*, *Myl6, and Rnd3*), oxidative phosphorylation and ATP metabolism (*Atp5b, Cox8a, and Ndufa7*), ribosome biogenesis (*Rpl13a),* water channel (*Aqp3*), and genes specific to embryo-linage (*Krt18*, *Ctsh, and Lgals1*). As shown in **Fig. 4A,** qRT-PCR results showed that all of these genes followed the same trend in terms of expression pattern, as with the RNA-seq results. Significant differences were validated for the expression levels of *Actb, Atp5b, Ctsh, Aqp3, Rpl13a, Krt18, Lgals1*, and *Rnd3* (P<0.05). In addition, the relative expression levels derived from RNA-seq were significantly and positively correlated to the expression levels obtained by qRT-PCR (R^2= 0.928, P<0.0001) **(Fig.4B)**.

Common gene identification between mouse hatching and human developmental progression

According to DEG analyses published in a previous study involving human embryo development (30), there were 4774 DEGs between E5 and E6 embryos (E5 *vs* E6), and 5253 DEGs between E6 and E7 embryos (E6 *vs* E7), when setting the cutoff criteria (adjusted P<0.05). By comparing these developmental progression-related DEGs and the hatching-related DEGs, we identified 40 common DEGs between human E5 *vs* E6 and mouse hatching DEGs, and 63 common DEGs between human E6 *vs* E7 and mouse hatching DEGs. Furthermore, there were 22 common DEGs among the three datasets (**Fig.5A-1**). Therefore, these 81 common DEGs were

assumed to be involved in the human embryo hatching process (Supplementary **Table 5**). In addition, when classifying DEGs into different lineages, more common DEGs were classified into the TE dataset, especially among the E7 *vs* E6 DEGs (77.8%) (**Fig. 5A-3**).

Downstream analysis suggested that these 81 common DEGs were enriched in 13 downstream pathways and 51 GO terms related to ribosome biogenesis and translation. In addition, the ATP metabolic process, oxidative phosphorylation, the citric acid (TCA) cycle, and the respiratory electron transport pathway, were also significantly enriched. Furthermore, downstream analysis showed that the actin cytoskeleton was potentially involved in human embryo hatching, as indicated by the enriched regulation of the actin cytoskeleton pathway and process (*ENAH, GSN, CFL1, PFN1, GRB2, RANBP1*, and *ACTB*) and the Rho cell motility signaling pathway (*GSN, CFL1,* and *PFN1*) (**Fig 6.B**). The PPI network and MCODE analysis revealed four densely connected gene networks (Supplementary **Fig 4**); module 3 (*KRT7, KRT8, CLDN4, KRT18*, and *AQP3*) and module 4 (*CFL1, PFN1, GSN, GRB2, TPI1, ALDOA, CTSH, LGALS1, PLA2G15, AKR1A1, MYL6,* and *SPRY1*) contained the same genes as those shown in the list of genes related to mouse embryo (module C and D).

Discussion

The phenomenon of blastocyst hatching still remains as one of the least studied aspects of early mammalian development, despite the fact that it is of paramount importance to implantation and pregnancy. Embryo hatching is likely to be a process that is regulated by multiple factors. The potential mechanisms revealed by previous studies of hatching includes embryo-derived proteases, the hydrostatic pressure

caused by blastocyst expansion, regulatory molecules, and the cell movement caused by cytoskeleton organization (1). However, most of the previous work on embryo hatching only investigated a select number of molecules; the exact molecular network and mechanisms involved still remain unclear.

Previous researchers aimed to investigate the transcriptome of hatching embryos by using cDNA microarrays and by profiling gene expression levels between pre-hatched and hatched *in vitro* cultured mouse blastocysts. However, the embryos in these two groups were collected at different development timepoints (E4.0 and E4.5) (8). Consequently, the results of this research could have been confounded since the mouse embryo undergoes a dynamic genetic change from E3.5 to E4.5 (44). This previous study revealed that some genes related to cell adhesion and migration, such as E-cadherin and NCAM, are highly expressed in E4.5 hatched blastocysts (8); However, these molecules were found to play a key role in blastocyst development and linage segregation in a subsequent study (44). Thus, it is important to control the study timepoint when investigating this transitory period of embryo hatching. In the present study, we used blastocysts from the same culture timepoint (E4.5) but with a different hatching status in order to avoid the confounding factor of different embryo stage.

According to our RNA-seq results, we found that the cytoskeleton system is highly involved in the blastocyst hatching process. The actin cytoskeleton system was first related to hatching in a previous pioneering study that showed that the expression of actin filaments rose significantly in TE cells and that the inhibition of actin polymerisation by cytochalasin B could completely impede the hatching of mouse embryos (45). Since then, two other molecules were demonstrated to be important in

embryo hatching: SK3 (small-conductance calcium-activated K^+ channels 3) and AmotL2 (angiomotin like 2). SK3 was demonstrated to regulate blastocyst hatching by controlling the formation of F-actin (7). In addition, AmotL2 (angiomotin like 2) was shown to be associated with E-cadherin and cytoskeletal actin; the tension generated by E-cadherin/AmotL2/actin filament play a crucial role in generating the forces required for hatching at the blastocyst stage in mouse embryos (46). In addition, these authors showed that the genetic inactivation of AmotL2 could block mouse blastocysts from hatching out of the ZP (7,46). Collectively, from previous data, it is implicated that TE cells are capable of briefly modifying their actin network to produce actomyosin contractility, thus forcing themself into shape during the hatching process. However, the involved mechanism pathway has yet to be determined (7,46). In this present study, we found that certain genes play a vital role in the hatching process, including *Actb, Myl6, Cfl1, Gsn, Pfn1, Arpc2, Rnd3, Grb2, Tuba1b, Tpi1, Shmt1, Ldlrap1, and Tmsb10;* these genes act by participating in the re-organisation of the cytoskeleton, especially with regards to actin polymerisation. Furthermore, IPA analysis suggested that the active regulation of actin-based motility by Rho, especially the RhoA signalling pathway, might represent the primary regulatory mechanism responsible for such changes. It is known that the RhoA pathway plays a crucial role in actomyosin contractility and does so by regulating cell motility, cell shape, and movement (47-49). Thus, for the first time, we have demonstrated how this network (*Arpc2, Cfl1, Gsn, Myl6, Pfn1, Setp9*, and *Rnd3*) might regulate actomyosin contractility *via* the RhoA signaling pathway, thus participating in the blastocyst hatching process in mice.

In addition, TEPs have been observed during the hatching in many species (5,14); however, the regulatory mechanisms underlying the biogenesis of TEPs remains unknown. Our RNA-seq results revealed that genes related to podosomes (*Actb, Gsn,*

and *Arpc2*) were upregulated in hatching embryos. Podosomes are the dynamic actin-rich cellular protrusions that degrade the extracellular matrix through local proteolysis; these structures consist of a dense core of F-actin and a ring structure of actin-associated proteins (50). Considering the F-actin enriched structure in hatching blastocysts and the ability of TEPs to degrade the matrix (51), it follows that podosomes may be implicated in the biogenesis of TEPs during the hatching process. Further studies, including immunoassays and functional assays, are now needed to confirm this hypothesis.

ATP powers all the energy-dependent activity, such as the synthesis of protein, nucleic acids, and all other molecules that make up organisms. In addition, ATP fuels transport of molecules across the membrane, cell movement and cell division (52). During embryo development, ATP production rises dramatically at the blastocyst stage, which is predominantly derived by oxidative phosphorylation. The increased level of ATP production at the blastocyst stage is due to the activity of Na^+/K^+ ATPase, together with the increased requirement for the synthesis of numerous macromolecules, including DNA, RNA, protein, and lipid (53,54). In this study, genes related to ATP metabolism were found to be upregulated in hatching blastocysts (*Cox8a*, *Ndufa5, Cox7c, Uqcrh*, and many ATP synthase coding genes). This suggested that ATP metabolism was activated during the hatching process. It has been demonstrated that most of the cellular ATP is utilized in protein synthesis *via* tRNA aminoacylation and GTP regeneration (55). Consistently, we found that genes related to the biogenesis of ribosomes and translation were upregulated in hatching blastocysts. Therefore, the enhanced level of protein synthesis explains the high consumption of ATP during embryo hatching, at least in part. Furthermore, ATP plays a vital role in the dynamics of actin filaments in the cytoskeleton (56); the hydrolysis of ATP in actin filaments has

been shown to modulate specific properties of the filament (actin polymerization), thus providing a pivotal regulation role for the life cycle of actin (57). We also found that genes related to the positive regulation of the cytoskeleton, along with actin polymerization or depolymerization (*Tmsb10, Pfn1, Gsn, Cfl1, Enah, Arpc2, Grb2*, and *Sema5a*) were upregulated in hatching blastocysts. Thus, the enhanced activity of actin cytoskeleton during embryo hatching may also require high production of ATP.

Another important factor that consumes ATP is the activity of Na^+/K^+ ATPase, an enzyme that is critical during blastocyst formation (58). Nevertheless, we observed no differences in the Na^+/K^+ ATPase subunit coding genes (*Atp1a1, Atp1a3, Atp1b1, Atp1b2, and Atp1b3)* when compared between hatching and expanded blastocysts; this suggested that the abundance of Na^+/K^+ ATPase (at lease at mRNA level) might be not crucial for the hatching process. This hypothesis agreed with previous findings, that once the blastocyst is fully expanded, Na^+/K^+ ATPase activity decreases to basal levels, and ATP is no longer required for Na^+/K^+ ATPase (59).

Interestingly, our RNA-seq results identified an upregulated network of *Aqp3*, *Cldn4*, *Krt18, Krt8*, and *Krt7*. By participating in cell-cell junctions, water channel activity, and water homeostasis, these genes might be heavily involved in blastocyst hatching. It is known that aquaporins (AQP3/7) mediate water movement across the TE and that the expression of Aqp3 is influenced by osmotic challenges in mammalian cells (11,60). Furthermore, a previous study demonstrated the establishment of tight junction seals (TJP1/2 (Tight junction protein 1/2)) and CLDN4/6 (Claudin-4/6)) can block the leakage of water during blastocyst formation (11,61,62). Recently, it has also been found that all TE cells are covered with a dense keratin filament network during the blastocyst stage (63). Together with the data revealed in the present study, these

findings suggest that the elevated osmotic potential of an expanded blastocyst can lead to an increase in the expression of *Aqp3*, thus further facilitating the flow of water into the cavity. Combined with the reinforcement of cell-cell junctions by tight junctions (*Cldn4)*, and keratin filaments (*Krt18, Krt8,* and *Krt7*), the further increase of hydrostatic pressure created by the expanded blastocysts will inevitably lead to embryo hatching. However, functional studies are now needed to investigate the importance of AQP3, CLDN4, and keratin filament, in the blastocyst hatching process.

Gene analysis of embryo linage showed that 16 TE-specific genes were all upregulated (i.e., *Ctsh* (cathepsin H), and *Lgals1* (lectin, galactose binding, soluble 1)), thus suggesting TE cells might be promoted in hatching blastocysts. This result is consistent with previous findings in that enhanced TE development could lead to the advancement of embryo hatching (64). Previously, the role of cathepsin P and L in hatching was investigated in the golden Syrian hamster (14). Our RNA-seq dataset demonstrated the increased expression of *Ctsh* in hatching blastocyst but not cathepsin B and L (*Ctsb* and *Ctsl*). This result indicated the potential role of cathepsin H in mouse embryo hatching. In addition, it was demonstrated that advanced TE development improves blastocyst attachment and trophoblast outgrowth *in vitro*; this may be related to the embryonic secretion of cathepsins and galectins (65-68). Consistently, the RNA-seq data demonstrated the increased levels of *Ctsh* and *Lgals1* in hatching blastocysts, this providing a clue that cathepsin H and galectin-1 may contribute to the hatching process or the preparatory process for early implantation in mice.

Furthermore, for the first time, we revealed novel upstream regulators during the hatching process by applying IPA software. Several transcription regulators were

predicted to be activated in the hatching embryos. MYC/MYCN have been implicated in the transcriptional regulation of several thousand genes involved in a wide range of cellular process, including cell-cycle control, metabolism, signal transduction, self-renewal, the maintenance of pluripotency, and the control of cell fate decision. MYC clearly plays an essential role in the process that supports early embryonic development and pluripotent stem cell biology (69). In this study, we found that MYC/MYCN may also play a role in the embryo hatching process. In addition, NFE2L2 (known as NRF2) is a transcription factor that regulates the cellular defense response and has been implicated to participate in early mouse embryo development by affecting embryo cleavage and blastocyst formation (70). In this study, NFE2L2 was also predicted to be active in hatching embryos by regulating downstream molecules (*Akr1a1, Aldoa, Calm1, Cox4i1, Dad1, Ppib, Rpl18, Rps16,* and *Tpi1*).

Growth factors, including AGT (Angiotensin) and FGF2 (beta fibroblast growth factor), were also found to positively regulate the hatching process in mouse embryos. Upstream analysis showed that the activation of AGT might upregulate the expression levels of *Calm1, Cox4i1, Cox7a2, Cox7c, Hmgb2, Pfn1, and Tbx20*. This could explain the previous finding that the supplementation of medium with angiotensin (Ang II) could improve embryonic development and increase the hatching rate (71). FGF2, a uterine factor, was previously found to be important for blastocyst formation (72,73). It has been reported that FGF2/ FGFR2 may act in an autocrine manner and activate the downstream PKC/p39 pathway during the formation of expanded mouse blastocysts (74). Interestingly, our results further showed that FGF2, by targeting *Aqp3, Tmsb10, Spry1, and Dad1,* could also positively influence the hatching process.

Apart from growth factors, the regulation of key hormones, including estrogen, PRL (prolactin), and insulin, is also important for embryo hatching. Estrogen-related receptors (ERRs) are nuclear receptors; there are three types of these receptors: Esrra, Esrrb, and Esrrg (75). Studies involving knockout animals have revealed that ERRs play a role in the regulation of metabolism and energy homeostasis. In particular, Esrrb was demonstrated to be important in embryo development by sustaining pluripotency and reprogramming (75). The present study, for the first time, showed that Esrra, the gene most closely related to the estrogen receptor, was predicted to be triggered in the hatching blastocyst, by actively regulating *Aldoa, Atp5f1b, Atp5mc1, Atp5pb, Atp5po, Cox7c, Cox8a,* and *Ndufa5*. Furthermore, the estrogen receptor was also found to actively regulate the expression of *Cldn4, Krt18, Krt7, Krt8,* and *Tmeff1.* These findings are consistent with a previous study in which the treatment of hamster embryos with an estrogen receptor-α (ER-α) antagonist induced a profound reversible inhibition of blastocyst hatching, thus indicating the definitive regulatory role of ER-α in the hatching process (5). Collectively, the available data suggests that estrogen, by activating Esrra and the estrogen receptor, can promote the hatching of mouse embryos *in vitro*.

PRL has been previously reported to be involved in blastocyst formation, as demonstrated by higher blastocyst rates when PRL was added to the culture medium (76). Furthermore, it was recently demonstrated that PRL signaling could stimulate the adhesion of blastocysts by promoting integrin-based focal adhesion and cytoskeletal organization during blastocyst outgrowth; it was also found that the addition of PRL to culture medium significantly increased human blastocyst outgrowth (77). Our results demonstrated that PRL could improve embryo hatching by targeting *Actb, Atp5h, Cox4i1, Tmsb10,* and *Tuba1b.* This may also explain the improved blastocyst

outgrowth shown in a previous study (77). Similarly, insulin also plays a role in embryo development, since the *in vitro* development of mouse embryos could be enhanced by the addition of insulin to the culture medium (78). Furthermore, it was shown that the proportion of hatching embryos was increased when insulin was added to the culture medium (78). In line with these findings, the analysis of upstream regulators in the present study suggested that insulin can positively regulate embryo hatching by targeting *Actb, Aldoa, Atp5mc1, Cox7c, Mdh2, Ppib,* and *Rpl32.* Previously, Igf1r signaling was demonstrated to be indispensable for preimplantation mouse embryo development (79). Studies in mouse and feline embryos further suggested that IGF-1 supplementation in culture medium improved the *in vitro* preimplantation development of embryos (80,81). Interestingly, we found that IGF1R, by regulating some genes related to metabolism, may also improve the embryo hatching process.

The RNA-seq and upstream analysis described herein demonstrate the connection between maternal hormones, growth factors, and the embryo hatching process. These findings could explain the observed difference between *in vitro* and *in vivo* hatching rates (82). In addition, this study identified potential genes for future investigations; studying these genes may enhance our understanding of how this upstream regulator is involved in embryo development, hatching and implantation. Furthermore, these hormones and growth factors might represent promising supplementation for *in vitro* culture media to enhance the efficiency of embryo hatching, although future studies now need to verify these findings in human embryos.

Many studies have investigated the human embryo linage and differentiation mechanisms using NGS techniques; however, there are no human data available that can allow us to profile transcriptomic regulation in the hatching process. It is known

that the human blastocyst hatches out from the ZP between E5 to E7 (1); thus, in this study, we compared our hatching-related gene dataset with a human developmental progression-related (from E5 to E7) gene dataset. We identified 81 common DEGs that may be involved in human embryo hatching. In addition, we also found that these DEGs were mainly TE-specific genes; thus, suggested that the TE is the main cell type participating in the hatching process in human embryos. The downstream analysis of these common DEGs suggested that ribosome biogenesis and translation, ATP metabolism, and regulation of the cytoskeleton, are important processes in the hatching of human blastocysts. In addition, according to these common DEGs, we also identified similar vital gene networks that may be important in the hatching process for human embryos, including genes related to water homeostasis and cell-cell junctions *(AQP3, CLDN4, KRT18, KRT8*, and *KRT7*) and genes related to the regulation of the cytoskeleton (e.g., *CFL1, PFN1, MYL6*, and *GRB2*). This study, for the first time, identified a suite of potential genes that participate in the hatching process of human embryos; however, future studies with donated human embryos are now necessary to study the specific role of these genes and investigate the underlying mechanisms.

Conclusion

This is the first study to characterise the transcriptome of mouse blastocysts during *in vitro* hatching by applying RNA-seq. Our findings indicated that the *in vitro* hatching of mouse embryos is an ATP-dependent process and that protein biosynthesis and the organization of the cytoskeleton play critical roles. The findings described herein also identified the novel genes, pathways, and upstream regulators, that are involved in the dynamic process of hatching, thus serving as the basis for further investigations targeted to the hatching mechanism. Furthermore, this study identifies candidate

molecules for future research relating to the human embryo hatching mechanism, which could potentially help to develop diagnostic methods or therapeutic targets to improve the competence of embryos for implantation.

F&S Science Clinical Quick Take

- RNA-seq data analysis revealed that the *in vitro* hatching of mouse blastocysts is an ATP-dependent process, and that protein biosynthesis and the cytoskeleton organization are highly activated. In addition, RhoA signaling could also be a vital regulatory pathway that participates in the hatching process.
- Key networks of genes involved in the regulation of cytoskeleton organization (e.g., *Actb, Myl6, Cfl1, Gsn, and Rnd3*) and cell-cell junction and water homeostasis *(Krt8, Krt7, Krt18, Cldn4*, and *Aqp3)* are implicated in the mouse hatching process.
- Upstream regulators, such as transcription factors (MYC and MYCN), growth factors (AGT and FGF2), and endocrine factors (oestrogen receptor, Esrra, and PRL, insulin, and IGF1R), are predicted to regulate hatching in mouse embryos.
- We identified potential genes involved in the human embryo hatching process; these represent candidate molecules for further human research.

Acknowledgement

This research was funded by the Rosetrees Trust (Reference CM541) awarded to KC. YL received a scholarship from the Chinese Council. We also thank the Wellcome Centre for Human Genetics (funded by Wellcome Trust grant reference 203141/Z/16/Z) for the generation and initial processing of the sequencing data.

References

1. Shafei R, Syrkasheva A, Romanov AY, Makarova N, Dolgushina N, Semenova M. Blastocyst hatching in humans. Russ J Dev Biol 2017;48:5-15.

2. Green D. Three-dimensional structure of the zona pellucida. Rev Reprod 1997;2:147-56.

3. Hill JA. Maternal‐embryonic cross‐talk. Ann N Y Acad Sci 2001;943:17-25.

4. Paria B, Reese J, Das SK, Dey SK. Deciphering the cross-talk of implantation: advances and challenges. Science 2002;296:2185-88.

5. Seshagiri PB, Roy SS, Sireesha G, Rao RP. Cellular and molecular regulation of mammalian blastocyst hatching. J Reprod Immunol 2009;83:79-84.

6. Leonavicius K, Royer C, Preece C, Davies B, Biggins JS, Srinivas S. Mechanics of mouse blastocyst hatching revealed by a hydrogel-based microdeformation assay. Proceedings of the National Academy of Sciences 2018;115:10375-80.

7. Lu Y-C, Ding G-L, Yang J, Zhang Y-L, Shi S, Zhang R-J, et al. Small-conductance calcium-activated K+ channels 3 (SK3) regulate blastocyst hatching by control of intracellular calcium concentration. Hum Reprod 2012;27:1421-30.

8. Chen H-W, Chen JJ, Yu S-L, Li H-N, Yang P-C, Su C-M, et al. Transcriptome analysis in blastocyst hatching by cDNA microarray. Hum Reprod 2005;20:2492-501.

9. Edashige K, Sakamoto M, Kasai M. Expression of mRNAs of the aquaporin family in mouse oocytes and embryos. Cryobiology 2000;40:171-75.

10. Barcroft LC, Offenberg H, Thomsen P, Watson AJ. Aquaporin proteins in murine trophectoderm mediate transepithelial water movements during cavitation. Dev Biol 2003;256:342-54.

11. Watson A, Natale D, Barcroft L. Molecular regulation of blastocyst formation. Anim Reprod Sci 2004;82:583-92.

12. Moriwaki K, Tsukita S, Furuse M. Tight junctions containing claudin 4 and 6 are essential for blastocyst formation in preimplantation mouse embryos. Dev Biol 2007;312:509-22.

13. Eckert JJ, Fleming TP. Tight junction biogenesis during early development. Biochimica et Biophysica Acta (BBA)-Biomembranes 2008;1778:717-28.

14. Sireesha G, Mason RW, Hassanein M, Tonack S, Navarrete Santos A, Fischer B, et al. Role of cathepsins in blastocyst hatching in the golden hamster. MHR: Basic science of reproductive medicine 2008;14:337-46.

15. Sharma N, Liu S, Tang L, Irwin J, Meng G, Rancourt DE. Implantation Serine Proteinases heterodimerize and are critical in hatching and implantation. BMC Dev Biol 2006;6:1-11.

16. Martin K, Barlow D, Sargent I. Heparin-binding epidermal growth factor significantly improves human blastocyst development and hatching in serum-free medium. Human Reproduction (Oxford, England) 1998;13:1645-52.

17. O'Sullivan CM, Ungarian JL, Singh K, Liu S, Hance J, Rancourt DE. Uterine secretion of ISP1 & 2 tryptases is regulated by progesterone and estrogen during pregnancy and the endometrial cycle. Molecular Reproduction and Development: Incorporating Gamete Research 2004;69:252-59.

18. Cheon YP. Effects of EGF and PAF on the Hatching and Implantation of Peri-implantation Stage Embryos. Development and Reproduction 2010;14:27-34.

19. Seshagiri PB, Vani V, Madhulika P. Cytokines and blastocyst hatching. Am J Reprod Immunol 2016;75:208-17.

20. Gonzales DS, Boatman DE, Bavister BD. Kinematics of trophectoderm projections and locomotion in the peri‐implantation hamster blastocyst. Developmental dynamics: an official publication of the American Association of Anatomists 1996;205:435-44.

21. Nagalakshmi U, Waern K, Snyder M. RNA‐Seq: a method for comprehensive transcriptome analysis. Curr Protoc Mol Biol 2010;89:4.11. 1-4.11. 13.

22. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nature Reviews Genetics 2019;20:631-56.

23. Wang QT, Piotrowska K, Ciemerych MA, Milenkovic L, Scott MP, Davis RW, et al. A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. Dev Cell 2004;6:133-44.

24. Hamatani T, Daikoku T, Wang H, Matsumoto H, Carter MG, Ko MS, et al. Global gene expression analysis identifies molecular pathways distinguishing blastocyst dormancy and activation. Proceedings of the National Academy of Sciences 2004;101:10326-31.

25. Zhang P, Zucchelli M, Bruce S, Hambiliki F, Stavreus-Evers A, Levkov L, et al. Transcriptome profiling of human pre-implantation development. PLoS One 2009;4:e7844.

26. Mamo S, Bodo S, Kobolak J, Polgar Z, Tolgyesi G, Dinnyes A. Gene expression profiles of vitrified in vivo derived 8‐cell stage mouse embryos detected by high density oligonucleotide microarrays. Molecular Reproduction and Development: Incorporating Gamete Research 2006;73:1380-92.

27. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 2013;20:1131.

28. Xue Z, Huang K, Cai C, Cai L, Jiang C-y, Feng Y, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 2013;500:593-97.

29. Deng Q, Ramsköld D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 2014;343:193-96.

30. Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. Cell 2016;165:1012-26.

31. García-Jiménez M, Rink K, Mestres E, Vanrell I, Calderón G, Costa-Borges N. Evaluation of the impact of laser-assisted hatching techniques on the hatching process of mouse blastocysts using time-lapse microscopy. F&S Science 2021.

32. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc 2014;9:171.

33. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nature methods 2015;12:357-60.

34. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923-30.

35. Anon R. R Development Core Team R: A language and environment for statistical computing. 2007.

36. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010 Jan 1;26:139-40.

37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B 1995;57.1:289-300.

38. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, et al. GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics 2004 Dec 12;20:3710-5.

39. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012 May;16:284-7.

40. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res 2016:gkw937.

41. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.

42. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003;4:2.

43. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods 2001;25:402-08.

44. Sozen B, Can A, Demir N. Cell fate regulation during preimplantation development: a view of adhesion-linked molecular interactions. Dev Biol 2014;395:73-83.

45. Cheon YP, Gye MC, Kim C-H, Kang BM, Chang YS, Kim SR, et al. Role of actin filaments in the hatching process of mouse blastocyst. Zygote 1999;7:123-29.

46. Hildebrand S, Hultin S, Subramani A, Petropoulos S, Zhang Y, Cao X, et al. The E-cadherin/AmotL2 complex organizes actin filaments required for epithelial hexagonal packing and blastocyst hatching. Sci Rep 2017;7:1-17.

47. Hanna S, El-Sibai M. Signaling networks of Rho GTPases in cell motility. Cell Signal 2013;25:1955-61.

48. Murrell M, Oakes PW, Lenz M, Gardel ML. Forcing cells into shape: the mechanics of actomyosin contractility. Nature reviews Molecular cell biology 2015;16:486-98.

49. Wheeler AP, Ridley AJ. Why three Rho proteins? RhoA, RhoB, RhoC, and cell motility. Exp Cell Res 2004;301:43-49.

50. Linder S, Kopp P. Podosomes at a glance. J Cell Sci 2005;118:2079-82.

51. Salas-Vidal E, Lomeli H. Imaging filopodia dynamics in the mouse blastocyst. Dev Biol 2004;265:75-89.

52. Lodish H, Berk A, Kaiser CA, Kaiser C, Krieger M, Scott MP, et al. Molecular cell biology. Macmillan; 2008.

53. Houghton FD, Leese HJ. Metabolism and developmental competence of the preimplantation embryo. European Journal of Obstetrics & Gynecology and Reproductive Biology 2004;115:S92-S96.

54. Leese HJ. Metabolism of the preimplantation embryo: 40 years on. Reproduction 2012;143:417.

55. Pontes MH, Sevostyanova A, Groisman EA. When too much ATP is bad for protein synthesis. J Mol Biol 2015;427:2586-94.

56. Becker EW. The roles of ATP in the dynamics of the actin filaments of the cytoskeleton. 2006.

57. Katkar HH, Davtyan A, Durumeric AE, Hocky GM, Schramm AC, Enrique M, et al. Insights into the cooperative nature of ATP hydrolysis in actin filaments. Biophys J 2018;115:1589-602.

58. Kidder GM, Watson AJ, editors. Roles of Na, K-ATPase in early development and trophectoderm differentiation. Seminars in nephrology; 2005: Elsevier.

59. Houghton FD, Humpherson PG, Hawkhead JA, Hall CJ, Leese HJ. Na+, K+, ATPase activity in the human and bovine preimplantation embryo. Dev Biol 2003;263:360-66.

60. Sugiyama Y, Ota Y, Hara M, Inoue S. Osmotic stress up-regulates aquaporin-3 gene expression in cultured human keratinocytes. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 2001;1522:82-88.

61. Sheth B, Nowak RL, Anderson R, Kwong WY, Papenbrock T, Fleming TP. Tight junction protein ZO-2 expression and relative function of ZO-1 and ZO-2 during mouse blastocyst formation. Exp Cell Res 2008;314:3356-68.

62. Xiong Y, Tan Y-J, Xiong Y-M, Huang Y-T, Hu X-L, Lu Y-C, et al. Expression of aquaporins in human embryos and potential role of AQP3 and AQP7 in preimplantation mouse embryo development. Cell Physiol Biochem 2013;31:649-58.

63. Lim HYG, Alvarez YD, Gasnier M, Wang Y, Tetlak P, Bissiere S, et al. Keratins are asymmetrically inherited fate determinants in the mammalian embryo. Nature 2020 2020/09/01;585:404-09.

64. Xu J-S, Cheung T-M, Ting-Hon Chan S, Ho P-C, Shu-Biu Yeung W. Temporal effect of human oviductal cell and its derived embryotrophic factors on mouse embryo development. Biol Reprod 2001;65:1481-88.

65. Afonso S, Romagnano L, Babiarz B. The expression and function of cystatin C and cathepsin B and cathepsin L during mouse embryo implantation and placentation. Development 1997;124:3415-25.

66. Afonso S, Romagnano L, Babiarz B. Expression of cathepsin proteinases by mouse trophoblast in vivo and in vitro. Dev Dyn 1999;216:374-84.

67. Barrientos G, Freitag N, Tirado-Gonzalez I, Unverdorben L, Jeschke U, Thijssen VL, et al. Involvement of galectin-1 in reproduction: past, present and future. Hum Reprod Update 2014;20:175-93.

68. You J-L, Wang W, Tang M-Y, Ye Y-H, Liu A-X, Zhu Y-M. A potential role of galectin-1 in promoting mouse trophoblast stem cell differentiation. Mol Cell Endocrinol 2018;470:228-39.

69. Chappell J, Dalton S. Roles for MYC in the establishment and maintenance of pluripotency. Cold Spring Harb Perspect Med 2013;3:a014381.

70. Lin Y, Sui L-C, Wu R-H, FU H-Y, XU J-J, QIU X-H, et al. Nrf2 inhibition affects cell cycle progression during early mouse embryo development. Journal of Reproduction and Development 2018;64:49-55.

71. Naderi MM, Borjian Boroujeni S, Sarvari A, Heidari B, Akhondi MM, Zarnani AH, et al. The Effect of Angiotensin on the Quality of In Vitro Produced (IVP) Sheep Embryos and Expression of Na(+)/K(+)/ATPase. Avicenna J Med Biotechnol 2016 Jan-Mar;8:9-15.

72. Fields SD, Hansen PJ, Ealy AD. Fibroblast growth factor requirements for in vitro development of bovine embryos. Theriogenology 2011 May;75:1466-75.

73. Yang QE, Fields SD, Zhang K, Ozawa M, Johnson SE, Ealy AD. Fibroblast growth factor 2 promotes primitive endoderm development in bovine blastocyst outgrowths. Biol Reprod 2011 Nov;85:946-53.

74. Yang J, Zhang D, Yu Y, Zhang RJ, Hu XL, Huang HF, et al. Binding of FGF2 to FGFR2 in an autocrine mode in trophectoderm cells is indispensable for mouse blastocyst formation through PKC-p38 pathway. Cell Cycle 2015;14:3318-30.

75. Festuccia N, Owens N, Navarro P. Esrrb, an estrogen‐related receptor involved in early development, pluripotency, and reprogramming. FEBS Lett 2018;592:852-77.

76. Kiapekou E, Loutradis D, Mastorakos G, Bletsa R, Beretsos P, Zapanti E, et al. Effect of PRL on in vitro follicle growth, in vitro oocyte maturation, fertilization and early embryonic development in mice. Cloning Stem Cells 2009 Jun;11:293-300.

77. Ezoe K, Miki T, Ohata K, Fujiwara N, Yabuuchi A, Kobayashi T, et al. Prolactin receptor expression and its role in trophoblast outgrowth in human embryos. Reprod Biomed Online 2021;42:699-707.

78. Mihalik J, Rehak P, Koppel J. The influence of insulin on the in vitro development of mouse and bovine embryos. Physiol Res 2000;49:347-54.

79. Bedzhov I, Liszewska E, Kanzler B, Stemmler MP. Igf1r signaling is indispensable for preimplantation development and is activated via a novel function of E-cadherin. PLoS Genet 2012;8:e1002609.

80. Harvey MB, Kaye PL. Insulin ‐ like growth factor ‐ 1 stimulates growth of mouse preimplantation embryos in vitro. Mol Reprod Dev 1992;31:195-99.

81. Thongkittidilok C, Tharasanit T, Sananmuang T, Buarpung S, Techakumphu M. Insulin-like growth factor-1 (IGF-1) enhances developmental competence of cat embryos cultured singly by modulating the expression of its receptor (IGF-1R) and reducing developmental block. Growth Horm IGF Res 2014;24:76-82.

82. Mahdavinezhad F, Kazemi P, Fathalizadeh P, Sarmadi F, Sotoodeh L, Hashemi E, et al. In vitro versus In vivo: Development-, Apoptosis-, and Implantation- Related Gene Expression in Mouse Blastocyst. Iran J Biotechnol 2019 Jan;17:e2157.

Figure legends

Figure 1 Differentially expressed genes (DEGs) and the linage-specific genes

(A) A volcano plot for DEGs: the x-axis shows the fold change in gene expression between samples, and the y-axis shows the statistical significance of the differences. The significantly up- and downregulated genes are highlighted in red and blue, respectively (adjusted P value<0.05). The top 50 genes by fold change are marked by gene name. The gray points represent non-significant DEGs. (B) Heatmap of relative expression level of the DEGs between hatching and pre-hatching blastocysts from all samples (pre-hatching (green) and hatching (blue) blastocysts). (C) A venn diagram of the DEGs and the ICM or TE-specific genes. (D) Heatmap of the expression of embryos lineage-specific genes among the DEGs between hatching and pre-hatching blastocysts. For the heatmap plots, yellow represents elevated levels (red was the highest level), while blue represents the decreased levels. The intensities were normalized for each gene in each row; samples and genes were clustered according to their similarity in profile using hierarchical cluster analysis. The annotation on the top showing clustering of the samples.

Figure 2 The downstream functional analysis of the DEGs

(A) The annotated Gene Ontology (GO) biological process term network of DEGs: the enriched terms are represented as separately nodes and organized into a network with edges connecting overlapping gene sets. The node sizes represent the gene numbers, and the term enrichment significance is represented by the colour (adjusted *P* value). (B) The top-ranked canonical pathways and (C) upstream regulators generated in Ingenuity Pathway Analysis (IPA). For each enriched pathway and regulators, the enrichment significance is plotted by a bar (-log(*P*-value)), and the predicted activity is presented in different colours (activated (orange), inhibited (blue), no predicted activity (gray), and Z-score=0 (white)). The Z-score of each predicted pathway is plotted by a red dot (x-axis above).

Figure 3 MCODE (molecular complex detection) analysis based on the PPI (protein-protein interaction) network

The nodes modules generated from PPI network using MCODE. The red nodes represent the up-regulated DEGs, the blue nodes represent the down-regulated DEGs, and the saturation of nodes indicates the absolute value of logFC (fold change); lines represent the interaction relationship between nodes. Module (A) contained 103 nodes, and 1805 edges; module (B) had 35 nodes and 176 edges; module (C) had 12 nodes and 16 edges; and module (D) contained 5 nodes, and 6 edges.

Figure 4 Validation of DEGs generated by RNA-sequencing using qRT-PCR

(A) The mRNA expression of 12 genes was detected using qRT-PCR. The amount of mRNA was quantified by measuring the cycle threshold (Ct) value. Relative expression of each transcript was normalized to mean Ct value of two housekeeping gens, and the ΔCt values of each sample were normalized to the mean ΔCt values of all pre-hatching samples (ΔΔCt). Data presented as mean ± SD, and statistical analysis was carried out using the unpaired t-test to compare pre-hatching and hatching groups (**P*<0.05). (B) The correlation between RNA-seq and qRT-PCR gene expression (R^2=0.93, *P*<0.0001)

Figure 5 The common DEGs between mouse hatching and human developmental progression-related gene lists and the downstream analysis

(A) The common DEGs representation through the Venn diagrams: the common DEGs between mouse embryo hatching-related and the human development progression-related gene lists (A-1), the common DEGs between mouse embryo hatching-related and human lineage-related gene lists in E6 *vs* E5 (A-2) and E7 *vs* E5 (A-3). (B) the downstream analysis of these common DEGs, including pathway analysis (B-1) and the GO term analysis (B-2).

Supplementary:

Figure S.1 Representative images of mouse blastocysts for RNA-seq experiment

Representative images for the frozen-thawed 8-cell embryo (E2.5), the early blastocyst at 24h of culture (E3.5) (400× magnification), and the pre-hatching and hatching blastocyst at 48h of culture (E4.5) (200× magnification) used in RNA-sequencing experiment (scale bar=50 μm).

Figure S.2 The advance quality control of RNA-seq data

The RNA-seq data was assessed from the read assignment by category (A, B), and the spearman correlation (C, D). The two outliners were marked in (A) and (C). For the name of each sample, the first and second number represents the embryo batch, and embryo number.

Figure S.3 The GO term analysis of cell components (CC) and molecular function (MF)

The enriched CC terms and MF terms are represented as separately nodes and organized into a network with edges connecting overlapping gene sets. The node sizes represent the gene numbers, and the term enrichment significance is represented by the colour (adjusted *P* value).

Figure S.4 The PPI network of mouse hatching-related genes (A), and the common DEGs between mouse hatching and human developmental progression-related gene lists (B)

(A) The red nodes represent the up-regulated DEGs, the blue nodes represent the down-regulated DEGs, and the saturation of nodes indicates the absolute value of logFC (fold change); lines represent the interaction relationship between nodes. (B) The PPI network of the common DEGs (B-1), and the node modules generated by MCODE analysis (B-2). Module 1 contained 35 nodes, and 284 edges; module 2 had 9 nodes and 19 edges; module 3 had 5 nodes and 7 edges; and module 4 contained 12 nodes, and 16 edges.