Blastocoel fluid RNA predicts pregnancy outcome in assisted reproduction

Authors

Caroline Frisendahl[1,2†], Kristina Gemzell-Danielsson[1,2†*] Karthik Nair[1], Forrest Koch[3], Judith Menezes[2], Mona Sheiki[4], Karin Dahlgren[4], Wesam Alnahal[4], Karin Lindgren[5], Therese Bohlin[6], Jan Olofsson[1,2], Virpi Töhönen[7], Fatemeh Vafaee[3], Parameswaran Grace Lalitkumar[1,2#*], Nageswara Rao Boggavarapu[1#], Omid R. Faridani[8,1#*]

* Correspondence: omid.faridani@ki.se; kristina.gemzell@ki.se; lalit.kumar@ki.se
† Shared first authors
# Shared last authors

Affiliations

1. WHO Collaborating Centre, Division of Neonatology, Obstetrics, Gynecology, and Reproductive Health, Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

2. Division of Gynecology and Reproductive Medicine, Karolinska University Hospital, Stockholm, Sweden

3. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia

4. Stockholm IVF, Stockholm, Sweden.

5. Reproduction center, Uppsala University Hospital, Uppsala, Sweden

6. Division of reproductive medicin, Örebro University Hospital, Örebro, Sweden

7. Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

8. School of Biomedical Sciences, University of New South Wales, Sydney, Australia

Abstract

Background and objective: Assessing embryo quality objectively and with greater precision remains a key challenge in assisted reproductive technologies (ART). RNA molecules excreted by the embryo are detectable in embryo spent media (ESM) and blastocoel fluid (BF). Their potential for predicting pregnancy outcome is however unknown. The aim of the study was to determine whether embryo excreted RNA can serve as biomarkers to predict pregnancy outcome in ART.

Material and methods: the primary outcome of this multicenter study was to develop a prediction model for implantation outcome based on mRNA molecules in ESM and BF. We included 255 couples undergoing ART with single embryo transfers between 2018 and 2019 in four IVF clinics in Sweden. ESM and BF were collected from individually cultured embryos. RNA extraction was performed using our novel protocol developed for ESM and BF. To identify the mRNA signature, each sample underwent mRNA sequencing (Smart-seq3), followed by differential gene expression analysis. A machine learning (ML) approach was applied with the mRNA genes in BF as input data to create a prediction model for pregnancy results measured as a positive or negative urine human chorionic gonadotropin (hCG) test result 14 days after embryo transfer.

Results: After pre-processing and quality controls, we included ESM from 105 cleavage stage and 230 blastocyst stage embryos and BF from 75 blastocyst stage embryos in the analysis. In total we identified 345 and 1132 mRNA genes in ESM from cleavage and blastocyst embryos respectively and 1157 mRNA genes in BF. There was a significant difference in RNA abundancy of 108 mRNA genes (p-value<0.05; fold change +2/-2) between hCG positive and negative embryos in BF but not in ESM. Consequently, the mRNA data from BF was selected as input data to train different ML models to predict pregnancy result after embryo transfer. The most optimal ML model combined L2 Regularized Logistic Regression with differentially abundant genes as feature and principal component analysis for preprocessing. The ML model predicted pregnancy result with a sensitivity of 83% and a specificity of 59% (AUC of 76% and accuracy of 69%). Testing of the model in an independent subset of 9 samples, not included in the training dataset, demonstrated a sensitivity of 60%, specificity of 75%, and accuracy of 67% in prediction of pregnancy result.

Conclusion: For the first time, we demonstrate that RNA molecules can be comprehensively mapped in individual ESM and BF samples using a highly sensitive RNA isolation and sequencing protocol. Moreover, mRNA in BF can be used to predict pregnancy outcome of individual embryos in ART using a ML approach. Our novel prediction model has the potential to enhance embryo selection, thereby improving the success rate of ART treatment. In the future, larger clinical trials are needed to validate the robustness of the method and compare it against standard care.

Keywords: Embryo spent media; Blastocoel fluid; cell-free RNA; Embryo quality; Machine learning; smart-seq3, mRNA sequencing

1. Introduction

Infertility affects about 12,5% of the global population in their reproductive age and has a great social, psychological and health economic impact (1-4). Advancements in assisted reproductive technologies (ART) have helped to combat this issue, but they have not fully solved it. About one in three couples remain infertile after several treatment attempts with ART (5). One of the key challenges in ART is to select the best quality embryo with the highest chance of resulting in a live birth upon transfer.

Morphological grading of embryos has been in practice over the past three decades as a method to assess embryo quality in ART. This method is, however, subjective, and frequently the assessment of embryos does not correlate with treatment outcome (6). With the aim of increasing objectivity and precision, AI algorithms based on morphokinetic and morphological parameters captured by time-lapse images, have been evaluated in many retrospective studies, but they have not yet demonstrated superiority over the conventional morphological scoring system in any prospective study (7, 8). Preimplantation genetic testing for aneuploidy (PGT-A) is frequently used as a complement when assessing embryos. This method is however time-consuming, expensive, requires highly skilled personnel and knowledge about the long-term impact of the invasive biopsy on the embryo is still limited (9-11). In addition, a recently published meta-analysis failed to demonstrate a benefit on live-birth rate or miscarriage rate when adding PGT-A to standard care (12).

Hence, there is an unmet need of objective, non-invasive biomarkers with improved precision to predict embryo quality in ART. In light of this, embryo spent culture media (ESM)—which contains embryo excreted molecules like RNA, DNA, metabolites, and proteins that may act as signaling molecules during implantation—has gained significant attention as a promising source of non-invasive biomarkers for assessing embryo quality (13-18). Initially research focused mainly on metabolites and proteins in ESM, but none of the suggested biomarkers has so far been translated into clinical care (18). Cell-free DNA is also currently studied in the context of developing a non-invasive PGT-A (16, 19). In recent years, small RNAs, especially microRNAs (miRNAs), found in ESM have been explored as potential biomarkers due to their importance in early embryo development and implantation (13, 14). Indeed, miRNA expression have shown to correlate with implantation outcome in IVF, but the results across studies have been inconsistent (20-26). This may indicate that miRNAs are less reliable as embryo quality biomarkers, but it could also reflect certain challenges linked to biomarker discovery in ESM. Given the tiny volume of ESM collected from only one embryo, the methods used to capture RNA expression need to be highly sensitive to distinguish the expression pattern from the background noise of the culture media itself (13, 27, 28). (13, 27, 28). Consequently, the sequencing data so far generated on miRNAs in ESM are mainly based on pooled samples, making it impossible to evaluate the ability of identified biomarkers to predict pregnancy outcome of individual embryos (20, 22, 23, 25, 26, 28-31).

In contrast to miRNAs, extracellular messenger RNAs (mRNAs) have not been explored as biomarkers for embryo implantation Considering the importance of protein coding RNAs during embryo development, this information may provide additional insights on the viability of the embryo not captured by only evaluating embryo morphology (32). A recently published study on mRNA in ESM from different stages of pre-implantation embryo development showed that the mRNA genes detected in ESM are embryo related and dynamically changing during embryo development (17). Apart from ESM, blastocoel fluid (BF) could also be a promising source of embryo quality markers. DNA content in BF has been suggested as a potential marker of embryo quality, the possibility of using cell-free mRNA in BF as markers for embryo quality has however not been explored (33, 34).

In this study, we therefore set out to explore the potential of using embryo excreted mRNA molecules in ESM and BF to predict pregnancy outcome in IVF by applying machine learning. To overcome the challenges mentioned above, we have developed an RNA extraction protocol tailored for low-molecular biofluids and adapted a highly sensitive mRNA sequencing method to enable the detection of mRNA molecules in individual samples of ESM and BF.

2. Material and Methods

2.1 Ethics

The study was approved by the Research Ethics Committee at Karolinska Institutet. Original approval (Dnr 00-134) and amendments (Dnr 2009/1615-32; Dnr 2012/1413-32, Dnr 2016/1290-32 and Dnr 2018/567-32). Collection of the included material has been performed after written informed consent from all study participants.

2.2 Study design

The study design is summarized as a flow-chart in Figure 1. This is a multicenter study with the aim to develop a prediction model for pregnancy outcome after embryo transfer in IVF based on the mRNA signature in ESM and BF. To achieve this, ESM have been collected from individually cultured cleavage and blastocyst stage embryos whereas BF has been collected from blastocyst embryos. To capture the mRNA signature in ESM and BF, we have subjected each individual sample to mRNA sequencing, followed by bioinformatic analyses. The resulting mRNA signatures have been used as input data to create a prediction model for pregnancy outcome using a machine learning approach.

2.3 Study population and IVF procedures

In total 255 couples undergoing IVF were included in the study from April 2018 to May 2019 from four IVF clinics in Sweden: IVF Stockholm, and the reproductive medicine units at Örebro University Hospital, Uppsala University Hospital and Karolinska University Hospital. All couples undergoing IVF or ICSI with single embryo transfers regardless of indication for the treatment were eligible to participate in the study. IVF or ICSI treatments utilizing donor sperm or donor oocytes have been excluded due to ethical concerns, as informed consent

cannot be obtained from the donor. Controlled ovarian stimulation was performed using either a short antagonist or long agonist Gonadotropin-Releasing Hormone (GnRH) protocol. Aspirated oocytes were fertilized *in-vitro* using either standard IVF or ICSI, according to the clinical routines and protocols used at each individual IVF clinic. Embryos were cultured in individual drops of either G-TL (Vitrolife, Sweden) or SAGE (CooperSurgical, Denmark) single-step culture media in a humidified incubator at 37°C with 6% CO_2 and 6% O_2.

2.4 Morphological assessment of the embryos

Morphological quality assessment of the embryos has been performed according to the Istanbul ESHRE consensus of 2011 (35). Cleavage stage embryos has been evaluated based on the number of blastomeres and the degree of fragmentation while blastocysts where scored based on the Gardner system that combine the developmental stage (blastocoel expansion and hatching status) with the quality of the inner cell mass (ICM) and trophoectoderm (TE). The developmental stage was scored from 1 to 6, where 6 correspond to a hatched blastocyst where the blastocyst has completely escaped from the zona pellucida. The ICM and TE were graded from A to C, where A correspond to the highest grade (top quality). To be considered for embryo transfer on day 3, cleavage stage embryos had to have 6-10 blastomeres and ≤ 20% fragmentation whereas blastocysts had to have a minimum Gardner score of 3BB. In rare cases (total 7 samples) blastocysts with a developmental stage less than 3 were transferred. Blastocysts that arrested in development or had a Grade C ICM and/or TE were defined as non-transferable (e.g. of morphological poor quality) and consequently discarded.

2.5 Collection of spent culture media and blastocoel fluid

ESM were collected from transfer-quality cleavage stage embryos on culture day 3 and from blastocyst embryos on culture day 5 or 6. We have also collected ESM on culture day 5 and 6 from non-transferable poor quality (PQ) embryos. Collection of ESM was performed immediately after fresh embryo transfer, vitrification or after discarding the embryo. A volume of 15-20ul of ESM was aspirated by carefully avoiding the overlaying oil and transferred to a sterile cryotube. BF was collected from blastocysts undergoing laser collapsing as a routine procedure before vitrification to avoid ice crystal formation. To collect BF, the embryo was transferred to a new drop of culture media where laser collapsing of the blastocoel cavity was performed resulting in a release of the blastocoel fluid into the culture media. A volume of 15-20ul of the BF containing culture media was collected in a sterile cryotube immediately after the collapsed embryo had been removed and vitrified. Culture media not exposed to an embryo but subjected to the same culture conditions and culture time, were collected as negative controls from each new batch of culture media used in the study. All samples were immediately stored in the vapour phase of liquid nitrogen until further processing.

2.6 Primary and secondary outcomes

The primary outcome of our study was to correlate mRNA molecules in ESM and BF with pregnancy result (positive or negative) 14 days after embryo transfer measured with a urine human chorionic gonadotropin (hCG) test. As secondary outcomes we evaluated the correlation of mRNA markers in ESM and BF between live-birth and negative pregnancy

outcomes (NPO). The NPO group included the following pregnancy outcomes defined according to the ESHRE guidelines (36): (1) A biochemical pregnancy was defined as a positive urine hCG test fourteen days after embryo transfer, without any visible gestational sac on ultrasound performed week 7 to 9 of gestation; (2) Miscarriage was defined as spontaneous loss of a previously verified intrauterine pregnancy (identified by a gestational sac or fetal heartbeat) on ultrasound before 22 weeks of gestation, with early miscarriage occurring before 13 weeks and late miscarriage after 13 weeks of gestation; and (3) a live birth was defined as a birth occurring after 22 weeks of gestation, where the newborn exhibits signs of life, regardless of whether the umbilical cord has been cut or if the placenta remains attached.

2.7 RNA extraction

The cryotube containing ESM or BF was briefly centrifuged at 2000rpm for 1 minute. The content of each sample was then transferred to an individual RNAse free 1.5ml Eppendorf tube (Sigma Aldrich, Sweden). Next, 100µl of Guanidium-based RLT plus buffer (Qiagen, Germany) was mixed with the sample followed by addition of 2ul of pellet paint to visualize the RNA pellet. To precipitate the RNA, 160µl of ethanol was subsequently added to the sample followed by vortexing and incubation for 48 hours at -20 degrees C. After incubation, the tubes were centrifuged at 16000g at 4^0C for 20min. The supernatant was then carefully discarded without disturbing the pink pellet. The RNA pellet was washed two times with 1ml of 80% fresh ethanol, air dried for 5 minutes and finally re-suspended in 7ul of RNase free water and stored at -80C until further processing.

2.8 Smart-seq3 mRNA libraries and sequencing

To detect cell-free mRNA molecules in ESM and BF, smart-seq3 (SS3) cDNA libraries were constructed from the extracted RNA of each sample. For this purpose, we have modified the published single-cell SS3 protocol to work optimally for low-input cell-free biofluids (37). Briefly, 3.8ul of extracted total RNA was used as input volume per sample. Water was included as negative controls. The mRNA was selectively hybridized using oligo-dT primers, followed by reverse transcription with Maxima H Minus reverse transcriptase enzyme (Thermo Scientific, Sweden) and by template switching oligo primers. Post reverse transcription, PCR amplification was performed with KAPA HiFi HotStart ReadyMix (Roche, Sweden) using 20 cycles of amplification. The amplified libraries were purified with Agencourt AMPure XP beads (Beckmann Coulter, USA) in a 0.7 to 1 ratio of beads to sample and finally eluted in 10µl of RNAse free water after purification. The purified libraries were quality checked using a High Sensitivity DNA Kit (Agilent technologies, USA) on an Agilent 2100 Bioanalyzer instrument (Agilent technologies, USA). The cDNA quantity was measured with a Qubit flex Fluorometer (Invitrogen, Singapore) using the Qubit 1X dsDNA High Sensitivity Kit (Invitrogen, Oregon, USA) as per the manufacturers protocol. Tagmentation of the amplified libraries was performed using Illumina Nextera XT kit (Illumina Inc, USA) as per the manufacturers recommendations with custom made Nextera XT dual index primers (IDT technologies, Germany). After tagmentation, each sample was purified with AMPure XP beads in a 1:1 ratio of beads to sample followed by quality and quantity evaluation using a bioanalyzer and Qubit flex Fluorometer respectively using the same reagents as described above. Finally, 5ng of cDNA from each

sample was pooled and sequenced on an Illumina Novoseq™ 6000 SP flow cell with a 100 base-pair single-end read setup (38).

2.9 Bioinformatic analyses

2.9.1 Processing of RNA sequencing data

The raw FASTQ files were quality-checked using FastQC (Version 0.11.9). Read alignment, annotation and quantification were performed according to the previously published zUMI pipeline (version 2.9.7) that was developed to handle UMI and sample specific barcodes and to allow for down sampling of hugely varying library sizes (39). The pipeline was run using the suggested miniconda environment. The demultiplexed reads generated for each sample were merged using the script 'merge_demultiplexed_fastq.R'. The merged fastq file was subjected to the zUMI pipeline. The samples were processed using pre-defined YAML configuration file against human reference GRCh38 and Gencode v44, with the following additional modifications to STAR aligner '--clip3pAdapterSeq TGTCTCTTATACACATCT --outFilterMismatchNmax 5 --outFilterMatchNmin 5'. In brief, the reads were filtered with a Phred score cut-off of >20 to filter out low quality reads. The remaining reads were mapped to the human genome 38 (hg 38) using STAR aligner (version 2.7.3a), followed by counting and UMI deduplication. The gene counts were then stored in an RDS file. Counts from exonic reads with UMIs were selected from the RDS file for downstream analysis. Tin.py module from RSeQC was used to generate Transcript Integrity Number (TIN) which was used to evaluate the RNA integrity of each sample. An mRNA molecule was considered as detected if at least 5 unique reads aligned to the gene.

2.9.2 Differential gene expression

To identify differentially abundant mRNAs in our data, we used the Bioconductor software DESeq2 (version 1.44.0). Differentially abundant genes (DAGs) with a p-value <0.05 and a fold change (FC) of ≤ -2 or ≥ 2 were considered statistically significant. The DESeq2 analyses was performed between the following groups: (1) hCG negative and positive embryos; (2) NPO and live-birth embryos, and (3) poor quality and hCG positive or live-birth embryos. ESM and BF samples were analyzed separately. BF and ESM samples from blastocysts were also analyzed based on maternal age where samples from women older than 34 years were analyzed separately from those of women younger than 35 years. For the ESM samples, cleavage and blastocyst embryos were analyzed separately. Further sub-analyses were also performed for blastocyst embryos according to developmental stage and culture day of transfer to understand their influence on the results. Sub-grouping of the BF samples according to the day of transfer and developmental stage deemed not feasible due to the small sample size.

2.9.3 Machine learning

Data were first divided into a training and testing set. The processed gene expression matrix was subset to include only protein coding RNA and counts of at least 1 across all samples.

Log-counts were obtained using the log1p transform. Differential gene expression analysis was performed using edgeR with a false discovery rate (FDR)<0.01 as cut-off for the DAGs. Leave-one-out cross validation (LOOCV) was then performed on the training set to identify an optimal model over all possible permutations of 3 parameters of interest: (1) model class (random forest (RF), L2 Regularized Logistic Regression (L2RLR), and L1 Regularized Logistic Regression (L1RLR)), (2) preprocessing (scaling and principal component analysis (PCA)), and (3) feature set (complete gene set, DAGs, morphological, complete gene set + morphological, and DAG + morphological) (40-42). Scaling, PCA, and differential gene expression analysis were all performed within the LOOCV iteration to avoid data leakage. The performance of the prediction models was evaluated using ROC curves and calculation of specificity, sensitivity, area under the curve (AUC), and accuracy (ACC) based on the confusion matrix generated. All analyses were performed using R version 4.4. The final selected parameters were: L2RLR, PCA, and DAGs. The model was retrained using all data from the training set and predictions were obtained for the testing set.

Results

mRNA content in embryo spent media and blastocoel fluid

In total we performed mRNA sequencing on 495 ESM and 148 BF samples. The RNA integrity in the sequenced samples, measured as the average TIN value across the cohort, was almost twice as high in the BF (TIN=1.31) compared to the ESM samples (TIN=0.70). We filtered out 88 ESM and 69 BF samples from vitrified embryos since they had not been transferred at the time of the analyses. A further 11 ESM and 4 BF samples were excluded because of poor read quality. In addition, 3 ESM samples transferred on culture day 4 were excluded. After the quality controls, ESM from 105 cleavage stage, 230 blastocyst stage and 58 PQ embryos that had underwent fresh or frozen transfers remained. The BF group finally included 75 samples from vitrified blastocyst embryos that had underwent frozen embryo transfers.

Maternal age, fertilization method and IVF cycle attempts were similar within and across the three main groups (Table 1). Fresh embryo transfers were more common in the cleavage stage embryos than blastocysts, reflecting current practice in Sweden. The proportion of fresh and frozen transfers in the hCG positive and negative cleavage stage embryos were however similar, thus will not influence the outcome of the DESeq2 analyses. In total we identified 345 and 1132 mRNA molecules in ESM from cleavage and blastocyst transfer quality embryos respectively. A majority (90%) of the mRNA genes detected in ESM from cleavage stage embryos was also detected in ESM from blastocyst embryos. In ESM from poor quality embryos we detected 270 mRNA genes. In BF, we identified 1157 mRNA genes and 72% of these mRNA genes were also detected in ESM from blastocyst embryos.

Weak association between mRNA in ESM and pregnancy outcome

To explore if the mRNA content in ESM correlats with the hCG outcome, we performed differential gene expression analysis. In ESM from cleavage stage embryos, we identified 5

DAGs between the hCG negative and positive embryos. Further comparison between NPO and live-birth embryos, revealed 5 DAGs out of which 3 genes overlapped with the genes identified in the hCG outcome comparison.

For blastocyst embryos transferred on culture day 5, only one gene was differentially abundant with higher abundance levels in the hCG negative group. Next, differential gene expression analysis was conducted separately for different stages of blastocyst development. Two genes were differentially abundant in developmental stage three blastocysts and three genes in stage four blastocysts. The NPO and live-birth analysis revealed no significant differences in mRNA content for blastocysts transferred on culture day 5. However, further subgroup analyses identified one, four, and four DAGs in blastocysts at developmental stages 3, 4, and 5, respectively.

mRNA content in ESM from PQ embryos were also compared to embryos with a positive hCG or a live birth outcome. We identified 15 DAGs between PQ and hCG-positive embryos and 18 DAGs between PQ and live birth embryos, suggesting that mRNA content in ESM exhibits greater diversity when comparing embryos with a positive pregnancy outcome to morphologically poor-quality embryos than when comparing them to morphologically good-quality embryos that resulted in a negative pregnancy outcome.

Strong association between mRNA in BF and pregnancy outcome

Next, we investigated if the mRNA content in BF is different in embryos with a negative compared to a positive pregnancy outcome. A significant increase in abundancy of 108 genes was identified in hCG negative compared to positive blastocyst embryos (Figure 2a). Similarly, 104 genes with an increased abundancy were detected between NPO and live birth embryos and out of these 52 genes overlapped with the hCG outcome comparison (Figure 2b).

Maternal age influences the mRNA signature in ESM and BF

Increasing maternal age is one of the main risk factors for aneuploidy in the embryo (43, 44). To understand the influence of age we divided the cohort into two groups with maternal age of 35 as the cut-off. In ESM, we identified three DAGs between hCG positive versus negative day 5 blastocyst embryos from women older than 34 in comparison to 0 DAGs in the group from women younger than 35. Interestingly, this pattern was more pronounced in BF, where 233 DAGs were identified between hCG-negative and positive blastocyst embryos in women older than 34 years, compared to only 10 DAGs in women younger than 35 years (Figure 3). A similar pattern, albeit not as distinct, was observed between BF from NPO and live birth as shown in the volcano plot of supplementary figure 1.

Successful prediction of pregnancy using machine learning based on mRNA in BF

The results of the differential gene expression analysis clearly demonstrated that the mRNA content in BF is more strongly associated with pregnancy outcome than the mRNA content in ESM. Consequently, we developed our machine learning model for prediction of pregnancy results based on the mRNA signature in BF. The performance of the different prediction

models is presented in table 2. The most optimal prediction model utilized L2RLR for machine learning with DAGs as input feature and PCA for pre-processing. The ML model predicted hCG outcome with a sensitivity of 83% and specificity of 59% (AUC of 0.76 and accuracy of 69%) as shown in the ROC curve of figure 4 and table 2. Testing of the model in a new subset of nine BF samples, which were not part of the training dataset, demonstrated a sensitivity of 60% and a specificity of 75% with an accuracy of 67% in predicting the hCG outcome of the transferred embryos.

Discussion

We show for the first time that RNA molecules can be extensively mapped in individual samples of ESM and BF by applying our novel RNA isolation and sequencing protocol. A clear difference in RNA content between pregnancy positive and negative embryos was observed in BF but not in ESM. A significant increase in abundancy of various mRNA genes was identified in BF from embryos with a negative pregnancy outcome. Our machine learning model could with a high accuracy predict pregnancy outcome of individual embryos based on the BF gene signature, which was successfully tested in a new subset of samples.

Our protocol for RNA extraction and optimized version of Smart-seq3 for mRNA sequencing have been able to detect numerous mRNA genes in both ESM and BF. Only one previous study, by Kirkegaard et al, has attempted to characterize the mRNA content in both ESM and BF, but failed to capture a mRNA signal above the background noise of the culture media, stressing the importance of the methodology used (27). ESM and BF are both low volume samples with RNA concentrations below the detection limit of conventional methods used in wet labs. The protocols in this study have been extensively optimized for the purpose of capturing the secretome of the embryo in ESM and BF. One other study has successfully captured mRNA content in ESM from preimplantation embryos using another mRNA sequencing protocol (17). However, in contrast to our study, the mRNA signal could not be linked to pregnancy outcome since ESM was collected from co-cultured embryos, limiting the clinical significance of their findings (17).

In our study, we successfully predicted hCG outcome after transfer with 69% accuracy and an AUC of 0,76 based on mRNA content in BF. Given the influence of endometrial factors on pregnancy outcomes, achieving a higher accuracy in an embryo-based prediction model may be challenging and could indicate an over-fitted model. No previous study has predicted pregnancy outcome based on molecular content in BF. Failure to amplify DNA content in BF has on the other hand been suggested to increase the likelihood of a live birth in a recently published prospective study (33). The lack of a proper control group in addition to a small sample size may however largely confound the data of that study, making it difficult to substantiate the relevance of their findings (33).

The potential of using ML to enhance embryo selection has gained increasing attention in recent years, with several ML models for embryo selection being published (7). The majority of the most robust studies have evaluated ML models trained on retrospective time-lapse images of transferred embryos, incorporating both morphological and morphokinetic

parameters (7). However, results from published studies on transferred blastocyst embryos have generally demonstrated inferior performance compared to our model in predicting pregnancy outcomes (45-48). For instance, a study by Berntsen et al. evaluated the KIDScore and IDAScore ML models, both of which exhibited lower precision than our model, with AUC values of 0.66 and 0.67, respectively (48). Similarly, a multivariate logistic regression model performed poorer than our model, with an AUC of 0.66. The same study did however also evaluate a random forest-based model that demonstrated comparable performance to ours, achieving an AUC of 0.74 (47). Notably, IDAScore is the only ML model that has been evaluated against manual morphological scoring in a double-blind randomized controlled trial, where it was found to be non-inferior to standard care in terms of clinical pregnancy rates (8). However, whether any of these models can significantly improve embryo selection remains unproven.

Laser collapsing of embryos is a routine procedure in many IVF labs to prevent ice crystal formation during vitrification and subsequent warming. In a clinical setting, our BF mRNA method could help assess embryos intended for frozen embryo transfers, particularly in clinics where laser collapsing is a standard practice. An objective test like this could both improve the precision in selection of embryos and reduce the number of unnecessary transfers of embryos that are unlikely to result in a live birth. This, in turn, could decrease the number of transfers required before achieving a successful pregnancy, ultimately improving patients' quality of life. Additionally, automating and standardizing embryo grading could optimize resource allocation and reduce costs for IVF clinics as well as patients and society. A key advantage of an mRNA-based test is its potential to be developed into a PCR-based panel. The widespread availability of commercial kits and primers for mRNAs makes its implementation in clinics both feasible and practical.

The findings from our study demonstrate that pregnancy outcome correlates to the mRNA signal detected in BF but not in ESM. There are several potential explanations for the absence of a clear signal in ESM. Compared to BF, the ESM content from cleavage and blastocyst stage embryos will contain a mixture of molecules excreted by the embryo during preimplantation development, which may dilute the signal associated with the implantation potential of the embryo. The embryo secretome in ESM may also get gradually degraded in the culture media. In contrast to miRNAs previously studied in this context, mRNAs are more susceptible to degradation. This hypothesis is supported by the lower TIN value of the ESM sample compared to BF indicating a lower integrity of the RNA. The ESM content may also be partially contaminated by sperm and cumulus cells. Addition of a washing step at the cleavage stage of embryo development could have eliminated any potential contamination at the blastocyst stage but this would have interfered with the clinical routine of using single-step media at the participating IVF centers. BF, on the other hand, is collected immediately after embryo collapsing in a fresh drop of culture media. Hence, the integrity of the RNA is presumably more intact and specific for the implantation stage. Potentially a larger sample size of ESM may have revealed more subtle changes in mRNA content.
Aneuploidy is a major cause of implantation failure, and first trimester miscarriages (49, 50). A recent study has investigated the fate of aneuploid cells in the preimplantation embryo (51).

The results of that study indicated that a majority of the aneuploid cells may be expelled into BF most likely as part of a self-correcting process, supporting the general increase in mRNA content in the embryos with a negative pregnancy outcome observed in our study (51). This is also in line with the stronger mRNA signal in BF that could be observed when embryos from women 35 years or older were analyzed separately. The increased abundancy of mRNA molecules could also be a mere reflection of a stressed embryo arresting in development resulting in apoptosis and release of cellular fragments into BF. Another likely function of the mRNAs is to serve as signaling molecules in the embryo-endometrial dialogue, which is essential for implantation. Supporting this hypothesis, a previous study has shown that poor-quality embryos exhibit an altered secretome, which, in vitro, appears to negatively impact the implantation process (52).

Several limitations should be considered when interpreting the data from this study. While the protocols we developed have high sensitivity for detecting mRNA molecules, the low molecular count data exhibit variability in expression patterns, which may reduce the method's overall sensitivity and reproducibility. This is a as proof-of-concept study that includes a limited sample size for training and testing the prediction model. Although a leave-one-out cross-validation strategy was used to mitigate overfitting of the machine learning model, larger clinical trials are necessary to further validate the reproducibiltiy of the method and compare it against morphological quality scoring of embryos. Additionally, the retrospective study design introduces a certain degree of selection bias that may affect the results since the analyzed embryos have been selected for transfer based on morphology and not the prediction of the mRNA profile in BF. A limitation that unfortunately cannot be avoided in the first step of developing a new model for embryo selection. Moreover, the lack of methods to reliably assess the receptivity of the endometrium to support embryo implantation remains a fundamental limitation in IVF treatment. Incorporating such parameters could potentially enhance the precision of this method. Similarly, information on chromosomal abnormalities using PGT-A might have improved the method's accuracy or facilitated embryo subgrouping. However, PGT-A is not permitted under Swedish law, making it unfeasible to include in this study. On the other hand, most of the clinical characteristics were well balanced between the positive and negative groups, minimizing the influence of those confounders on the outcome.

Conclusion

We show for the first time that RNA molecules can be extensively mapped in individual samples of ESM and BF by applying our novel RNA isolation and sequencing protocol. A clear difference in RNA content between pregnancy positive and negative embryos was observed in BF but not in ESM. A significant increase in secretion of mRNA genes was identified in BF from embryos with a negative pregnancy outcome. Our machine learning model could with a high accuracy predict pregnancy outcome of individual embryos based on the BF gene signature, which was successfully tested in a new subset of samples. A non-invasive technique if shown to be reliable in identifying good quality embryos would offer significant benefit to clinical practice and would save costs and resources. Larger clinical

trials are needed to validate the robustness of the method and compare it against standard care.

Acknowledgements
We want to acknowledge all the IVF patients that have contributed with invaluable material for this study and all the midwives, physicians and embryologists that have kindly assisted with the sample collection at each IVF clinic. The data processing was enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

Funding
This work was financially supported by the Swedish Research Council (2017-00932 and 2021-01042), joint grants from Region Stockholm and Karolinska Insitutet (ALF Medicine (954072, 973027), Research-AT 2020, FoUI-1001783).

Author contributions
KGD was responsible for overall study conduct conception, and funding. NRB contributed with the original research idea and conception. PGL conceptualized and led the study, contributed with insights on clinical embryology along with the original research idea. PGL and ORF did the feasibility studies. KGD, NRB, PGL, CF and ORF designed the study. CF, KGD and PGL obtained the regulatory permits. KGD, PGL, NRB, ORF and JM developed the sample collection protocol. KGD, CF, ORF, PGL and NRB planned the sample collection. CF and NRB coordinated the multi-center sample collection. WA, MS, KL, VT, JO, JM and TB were responsible on-site for the sample collection at each individual IVF clinic. NRB and ORF developed the RNA extraction protocol. CF optimized Smart-seq 3 for ESM and BF under supervision from NRB and ORF. NRB, ORF and CF designed the final sequencing experiments and NRB constructed the cDNA libraries. CF performed cleaning and quality controls of the clinical data. KN performed the bioinformatic analyses of the mRNA sequencing data. FK created the ML model with supervisonfrom FV and KN refined it. CF, PGL, ORF and NRB designed the data analysis, provided continuous scientific input and guidance during the data analyses. CF, KN, NRB, ORF, KGD and PGL interpreted the results. CF drafted the manuscript. All authors critically reviewed, edited and approved on the final version of the manuscript.

Competing interests
The authors have no competing interest to declare

References

1. WHO. Infertility prevalence estimates, 1990–2021. Geneva: World Health Organization; 2023. Licence: CC BY-NC-SA 3.0 IGO. 2023.
2. Chiware TM, Vermeulen N, Blondeel K, Farquharson R, Kiarie J, Lundin K, et al. IVF and other ART in low- and middle-income countries: a systematic landscape analysis. Human reproduction update. 2021;27(2):213-28.

3. Gerrits T, Kroes H, Russell S, van Rooij F. Breaking the silence around infertility: a scoping review of interventions addressing infertility-related gendered stigmatisation in low- and middle-income countries. Sex Reprod Health Matters. 2023;31(1):2134629.
4. Xie Y, Ren Y, Niu C, Zheng Y, Yu P, Li L. The impact of stigma on mental health and quality of life of infertile women: A systematic review. Front Psychol. 2022;13:1093459.
5. Smith A, Tilling K, Nelson SM, Lawlor DA. Live-Birth Rate Associated With Repeat In Vitro Fertilization Treatment Cycles. Jama. 2015;314(24):2654-62.
6. Lundin K, Ahlstrom A. Quality control and standardization of embryo morphology scoring and viability markers. Reproductive biomedicine online. 2015;31(4):459-71.
7. Salih M, Austin C, Warty RR, Tiktin C, Rolnik DL, Momeni M, et al. Embryo selection through artificial intelligence versus embryologists: a systematic review. Hum Reprod Open. 2023;2023(3):hoad031.
8. Illingworth PJ, Venetis C, Gardner DK, Nelson SM, Berntsen J, Larman MG, et al. Deep learning versus manual morphology-based embryo selection in IVF: a randomized, double-blind noninferiority trial. Nat Med. 2024.
9. Kokkali G, Coticchio G, Bronet F, Celebi C, Cimadomo D, Goossens V, et al. ESHRE PGT Consortium and SIG Embryology good practice recommendations for polar body and embryo biopsy for PGT. Hum Reprod Open. 2020;2020(3):hoaa020.
10. Neumann K, Sermon K, Bossuyt P, Goossens V, Geraedts J, Traeger-Synodinos J, et al. An economic analysis of preimplantation genetic testing for aneuploidy by polar body biopsy in advanced maternal age. BJOG : an international journal of obstetrics and gynaecology. 2020;127(6):710-8.
11. Alteri A, Cermisoni GC, Pozzoni M, Gaeta G, Cavoretto PI, Viganò P. Obstetric, neonatal, and child health outcomes following embryo biopsy for preimplantation genetic testing. Human reproduction update. 2023;29(3):291-306.
12. Cornelisse S, Zagers M, Kostova E, Fleischer K, van Wely M, Mastenbroek S. Preimplantation genetic testing for aneuploidies (abnormal number of chromosomes) in in vitro fertilisation. The Cochrane database of systematic reviews. 2020;9(9):Cd005291.
13. Hawke DC, Watson AJ, Betts DH. Extracellular vesicles, microRNA and the preimplantation embryo: non-invasive clues of embryo well-being. Reproductive biomedicine online. 2021;42(1):39-54.
14. Huang W, Chen ACH, Ng EHY, Yeung WSB, Lee YL. Non-Coding RNAs as Biomarkers for Embryo Quality and Pregnancy Outcomes: A Systematic Review and Meta-Analysis. Int J Mol Sci. 2023;24(6).
15. Poh QH, Rai A, Salamonsen LA, Greening DW. Omics insights into extracellular vesicles in embryo implantation and their therapeutic utility. Proteomics. 2023;23(6):e2200107.
16. Rubio C, Navarro-Sánchez L, García-Pascual CM, Ocali O, Cimadomo D, Venier W, et al. Multicenter prospective study of concordance between embryonic cell-free DNA and trophectoderm biopsies from 1301 human blastocysts. Am J Obstet Gynecol. 2020;223(5):751.e1-.e13.
17. Wu Q, Zhou Z, Yan Z, Connel M, Garzo G, Yeo A, et al. A temporal extracellular transcriptome atlas of human pre-implantation development. Cell Genom. 2024;4(1):100464.

18. Hernández-Vargas P, Muñoz M, Domínguez F. Identifying biomarkers for predicting successful embryo implantation: applying single to multi-OMICs to improve reproductive outcomes. Human reproduction update. 2020;26(2):264-301.
19. Sakkas D, Navarro-Sánchez L, Ardestani G, Barroso G, Bisioli C, Boynukalin K, et al. The impact of implementing a non-invasive preimplantation genetic testing for aneuploidies (niPGT-A) embryo culture protocol on embryo viability and clinical outcomes. Human reproduction (Oxford, England). 2024;39(9):1952-9.
20. Fang F, Li Z, Yu J, Long Y, Zhao Q, Ding X, et al. MicroRNAs secreted by human embryos could be potential biomarkers for clinical outcomes of assisted reproductive technology. J Adv Res. 2021;31:25-34.
21. Rosenbluth EM, Shelton DN, Wells LM, Sparks AE, Van Voorhis BJ. Human embryos secrete microRNAs into culture media--a potential biomarker for implantation. Fertility and sterility. 2014;101(5):1493-500.
22. Cuman C, Van Sinderen M, Gantier MP, Rainczuk K, Sorby K, Rombauts L, et al. Human Blastocyst Secreted microRNA Regulate Endometrial Epithelial Cell Adhesion. EBioMedicine. 2015;2(10):1528-35.
23. Borges E, Jr., Setti AS, Braga DP, Geraldo MV, Figueira RC, Iaconelli A, Jr. miR-142-3p as a biomarker of blastocyst implantation failure - A pilot study. JBRA Assist Reprod. 2016;20(4):200-5.
24. Capalbo A, Ubaldi FM, Cimadomo D, Noli L, Khalaf Y, Farcomeni A, et al. MicroRNAs in spent blastocyst culture medium are derived from trophectoderm cells and can be explored for human embryo reproductive competence assessment. Fertility and sterility. 2016;105(1):225-35.e1-3.
25. Abu-Halima M, Hausler S, Backes C, Fehlmann T, Staib C, Nestel S, et al. Micro-ribonucleic acids and extracellular vesicles repertoire in the spent culture media is altered in women undergoing In Vitro Fertilization. Scientific reports. 2017;7(1):13525.
26. Kamijo S, Hamatani T, Sasaki H, Suzuki H, Abe A, Inoue O, et al. MicroRNAs secreted by human preimplantation embryos and IVF outcome. Reprod Biol Endocrinol. 2022;20(1):130.
27. Kirkegaard K, Yan Y, Sørensen BS, Hardarson T, Hanson C, Ingerslev HJ, et al. Comprehensive analysis of soluble RNAs in human embryo culture media and blastocoel fluid. Journal of assisted reproduction and genetics. 2020;37(9):2199-209.
28. Russell SJ, Menezes K, Balakier H, Librach C. Comprehensive profiling of Small RNAs in human embryo-conditioned culture media by improved sequencing and quantitative PCR methods. Syst Biol Reprod Med. 2020;66(2):129-39.
29. Sánchez-Ribas I, Diaz-Gimeno P, Quiñonero A, Ojeda M, Larreategui Z, Ballesteros A, et al. NGS Analysis of Human Embryo Culture Media Reveals miRNAs of Extra Embryonic Origin. Reprod Sci. 2019;26(2):214-22.
30. Cimadomo D, Rienzi L, Giancani A, Alviggi E, Dusi L, Canipari R, et al. Definition and validation of a custom protocol to detect miRNAs in the spent media after blastocyst culture: searching for biomarkers of implantation. Human reproduction (Oxford, England). 2019;34(9):1746-61.
31. Abu-Halima M, Khaizaran ZA, Ayesh BM, Fischer U, Khaizaran SA, Al-Battah F, et al. MicroRNAs in combined spent culture media and sperm are associated with embryo quality and pregnancy outcome. Fertility and sterility. 2020;113(5):970-80.e2.

32. Petropoulos S, Edsgard 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(4):1012-26.
33. Gianaroli L, Perruzza D, Albanese C, Azzena S, Tabanelli C, Ferraretti AP, et al. Failure to detect DNA in blastocoel fluid is associated with a higher live birth rate in both PGT-A and conventional IVF/ICSI cycles. Human reproduction (Oxford, England). 2023;38(7):1268-76.
34. Magli MC, Albanese C, Crippa A, Tabanelli C, Ferraretti AP, Gianaroli L. Deoxyribonucleic acid detection in blastocoelic fluid: a new predictor of embryo ploidy and viable pregnancy. Fertility and sterility. 2019;111(1):77-85.
35. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Human reproduction (Oxford, England). 2011;26(6):1270-83.
36. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Human reproduction (Oxford, England). 2017;32(9):1786-801.
37. Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks GJ, Larsson AJM, et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nature biotechnology. 2020;38(6):708-14.
38. Wang L, Nie J, Sicotte H, Li Y, Eckel-Passow JE, Dasari S, et al. Measure transcript integrity using RNA-seq data. BMC Bioinformatics. 2016;17:58.
39. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. Gigascience. 2018;7(6).
40. Park MY, Hastie T. Penalized logistic regression for detecting gene interactions. Biostatistics. 2008;9(1):30-50.
41. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.
42. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1):1-22.
43. Gruhn JR, Zielinska AP, Shukla V, Blanshard R, Capalbo A, Cimadomo D, et al. Chromosome errors in human eggs shape natural fertility over reproductive life span. Science (New York, NY). 2019;365(6460):1466-9.
44. Charalambous C, Webster A, Schuh M. Aneuploidy in mammalian oocytes and the impact of maternal ageing. Nat Rev Mol Cell Biol. 2023;24(1):27-44.
45. Kragh MF, Rimestad J, Berntsen J, Karstoft H. Automatic grading of human blastocysts from time-lapse imaging. Comput Biol Med. 2019;115:103494.
46. Reignier A, Girard JM, Lammers J, Chtourou S, Lefebvre T, Barriere P, et al. Performance of Day 5 KIDScore™ morphokinetic prediction models of implantation and live birth after single blastocyst transfer. Journal of assisted reproduction and genetics. 2019;36(11):2279-85.
47. Blank C, Wildeboer RR, DeCroo I, Tilleman K, Weyers B, de Sutter P, et al. Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. Fertility and sterility. 2019;111(2):318-26.
48. Berntsen J, Rimestad J, Lassen JT, Tran D, Kragh MF. Robust and generalizable embryo selection based on artificial intelligence and time-lapse image sequences. PLoS One. 2022;17(2):e0262661.
49. Essers R, Lebedev IN, Kurg A, Fonova EA, Stevens SJC, Koeck RM, et al. Prevalence of chromosomal alterations in first-trimester spontaneous pregnancy loss. Nat Med. 2023;29(12):3233-42.

50. Peng G, Zhou Q, Chai H, Wen J, Zhao H, Taylor HS, et al. Estimation on risk of spontaneous abortions by genomic disorders from a meta-analysis of microarray results on large case series of pregnancy losses. Mol Genet Genomic Med. 2023;11(8):e2181.
51. Griffin DK, Brezina PR, Tobler K, Zhao Y, Silvestri G, McCoy RC, et al. The human embryonic genome is karyotypically complex, with chromosomally abnormal cells preferentially located away from the developing fetus. Human reproduction (Oxford, England). 2023;38(1):180-8.
52. Cuman C, Menkhorst EM, Rombauts LJ, Holden S, Webster D, Bilandzic M, et al. Preimplantation human blastocysts release factors that differentially alter human endometrial epithelial cell adhesion and gene expression relative to IVF success. Human reproduction (Oxford, England). 2013;28(5):1161-71.

Table 1. Clinical characteristics of the transferred embryos included in the mRNA analysis of ESM from cleavage and blastocyst stage embryos and BF from blastocyst stage embryos.

	ESM Cleavage stage		ESM Blastocyst stage		BF Blastocyst stage	
Parameter	hCG positive (n= 42)	hCG negative (n= 63)	hCG positive (n= 111)	hCG negative (n=119)	hCG positive (n=38)	hCG negative (n=37)
Maternal age (mean)	34	34	32	33	32	34
Fresh embryo transfer (%)	93%	90%	42%	49%	0%	0%
Frozen embryo transfer (%)	7%	10%	58%	51%	100%	100%
IVF (%)	62%	67%	66%	63%	68%	68%
ICSI (%)	38%	33%	34%	37%	32%	32%
Number of IVF cycle attempts (mean)	2	2	2	2	2	2
Culture day 5 at transfer/freezing (%)	---	---	92%	87%	89%	84%
Culture day 6 at freezing (%)	---	---	8%	13%	11%	16%
Biochemical pregnancy (%)	2%	---	7%	---	3%	---
Miscarriage (%)	14%	---	16%	---	18%	---
Ectopic pregnancy (%)	5%	---	0%	---	0%	---
Live birth (%)	79%	---	72%	---	68%	---

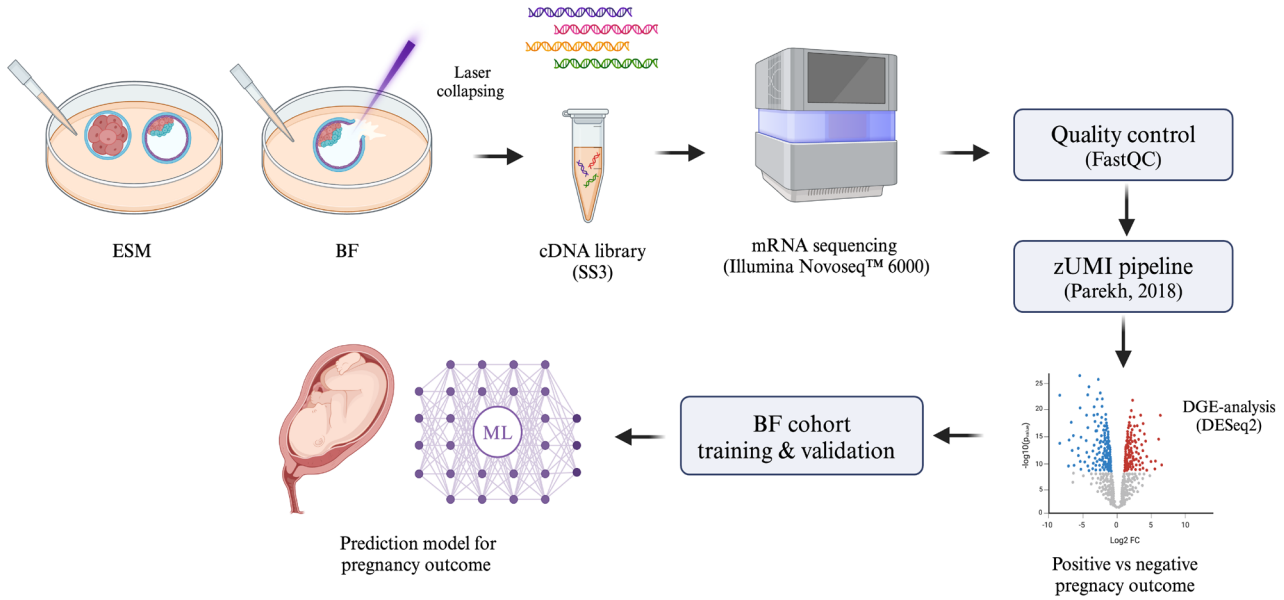
Abbreviations: ESM=embryo spent media; BF=blastocoel fluid; IVF=in-vitro fertilization; ICSI=intracytoplasmic sperm injection; hCG=human chorionic gonadotropin.

Table 2. Evaluation of ML models to predict hCG outcome in BF. Prediction of hCG outcome after embryo transfer was performed using different ML models trained on mRNA content in BF detected with mRNA sequencing.

Model class	Pre-processing	Input features	Specificity	Sensitivity	ACC	AUC
L2RLR	PCA	DAG	59%	83%	69%	0.76
L2RLR	PCA	All genes	59%	46%	53%	0.59
L2RLR	None	All genes	59%	46%	53%	0.62
L2RLR	None	morph+All genes	59%	46%	53%	0.61
L1RLR	PCA	DAG	56%	79%	66%	0.76
L1RLR	PCA	All genes	53%	54%	53%	0.55
L1RLR	None	DAG	59%	71%	64%	0.73
L1RLR	None	All genes	65%	42%	55%	0.56
L1RLR	None	morph+all genes	53%	21%	40%	0.37
L1RLR	None	morph+DAG	65%	67%	66%	0.75
L1RLR	None	morph	76%	29%	57%	0.48
RF	None	DAG	38%	83%	59%	0.58
RF	None	All genes	71%	21%	50%	0.45
RF	None	morph+DAG	44%	83%	60%	0.66
RF	None	morph+all genes	65%	21%	48%	0.45
RF	None	morph	76%	33%	62%	0.53

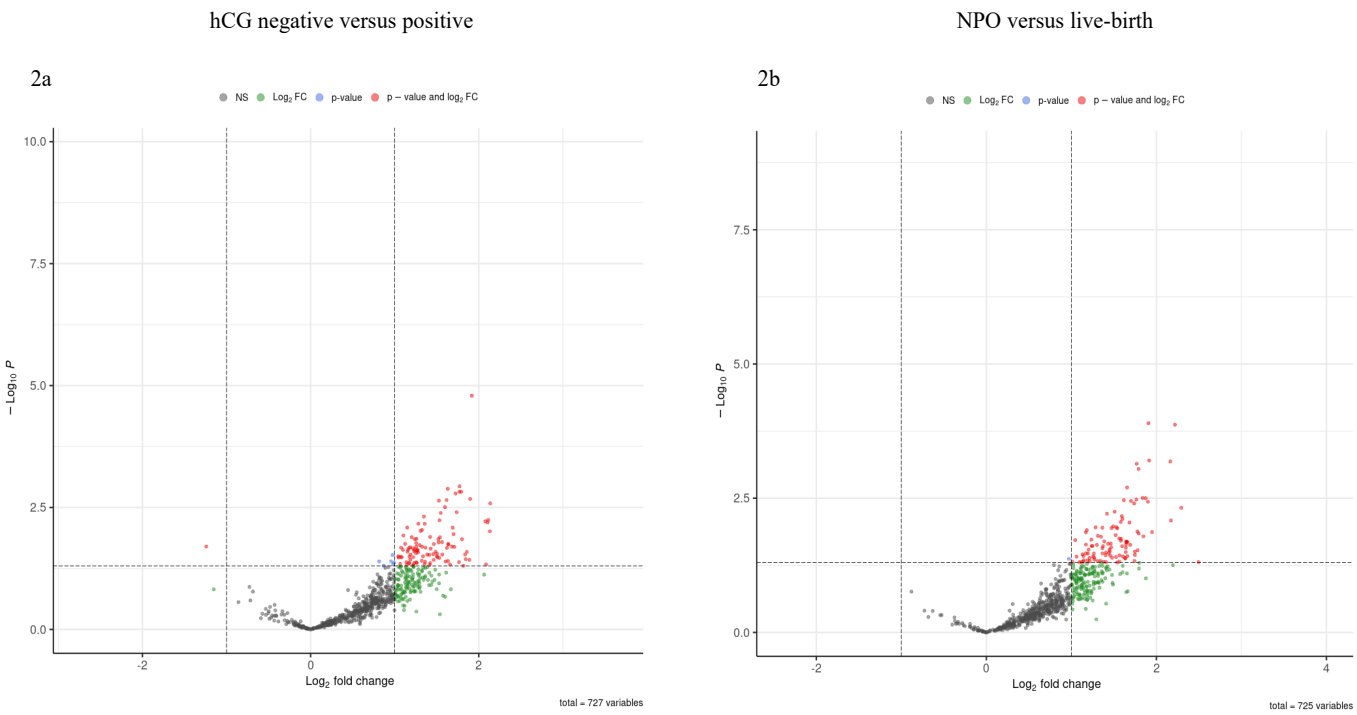
Abbreviations: ML=machine learning; BF=blastocoel fluid; L2RLR=L2 Regularized Logistic Regression, L1RLR=L1 Regularized Logistic Regression; RF=Random Forest; morph=morphology scores; DAG=differentially abundant genes with FDR<0.01; ACC=accuracy; AUC=area under the curve; hCG=human chorionic gonadotropin.

Figure 1. Flow-chart of the study design



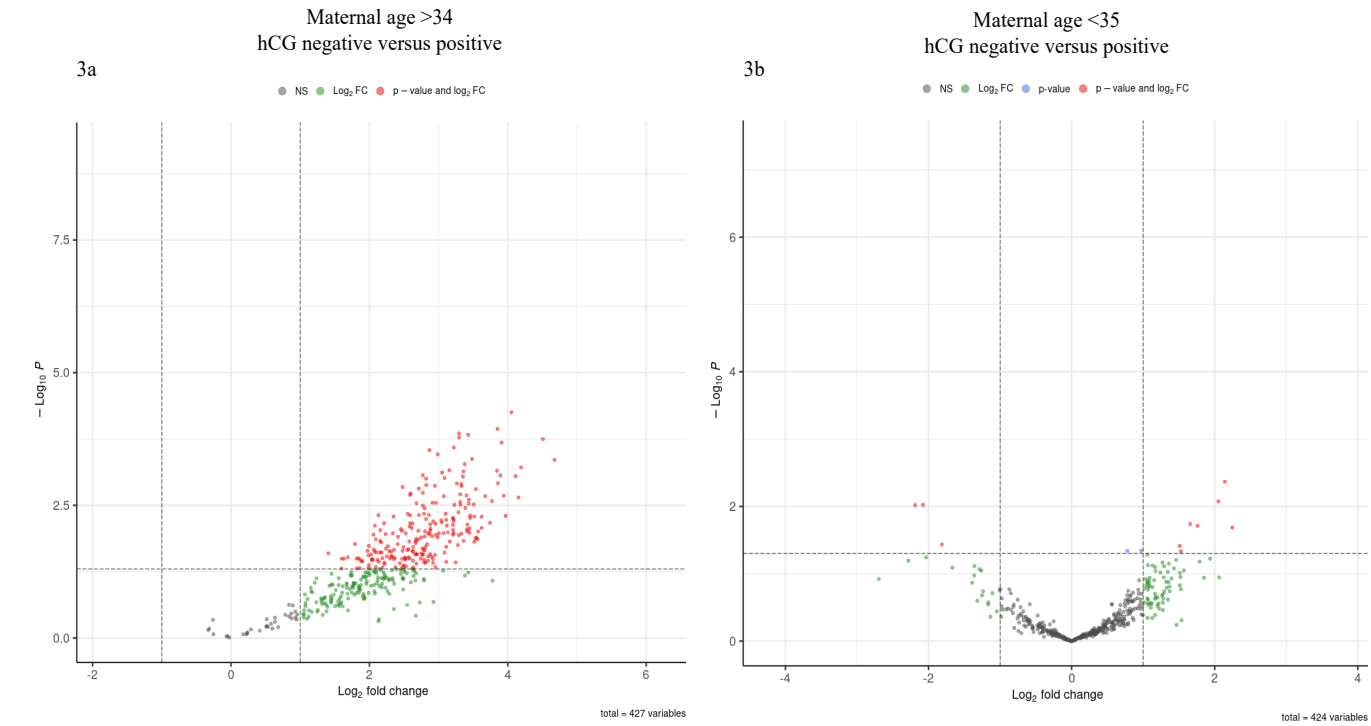
Abbreviations: ESM=embryos spent media; BF= blastocoel fluid; DGE=differential gene expression; ML=machine learning. Created in BioRender. Alkasalias, T. (2025) https://BioRender.com/u72a686

Figure 2. Differentially abundant mRNA in BF



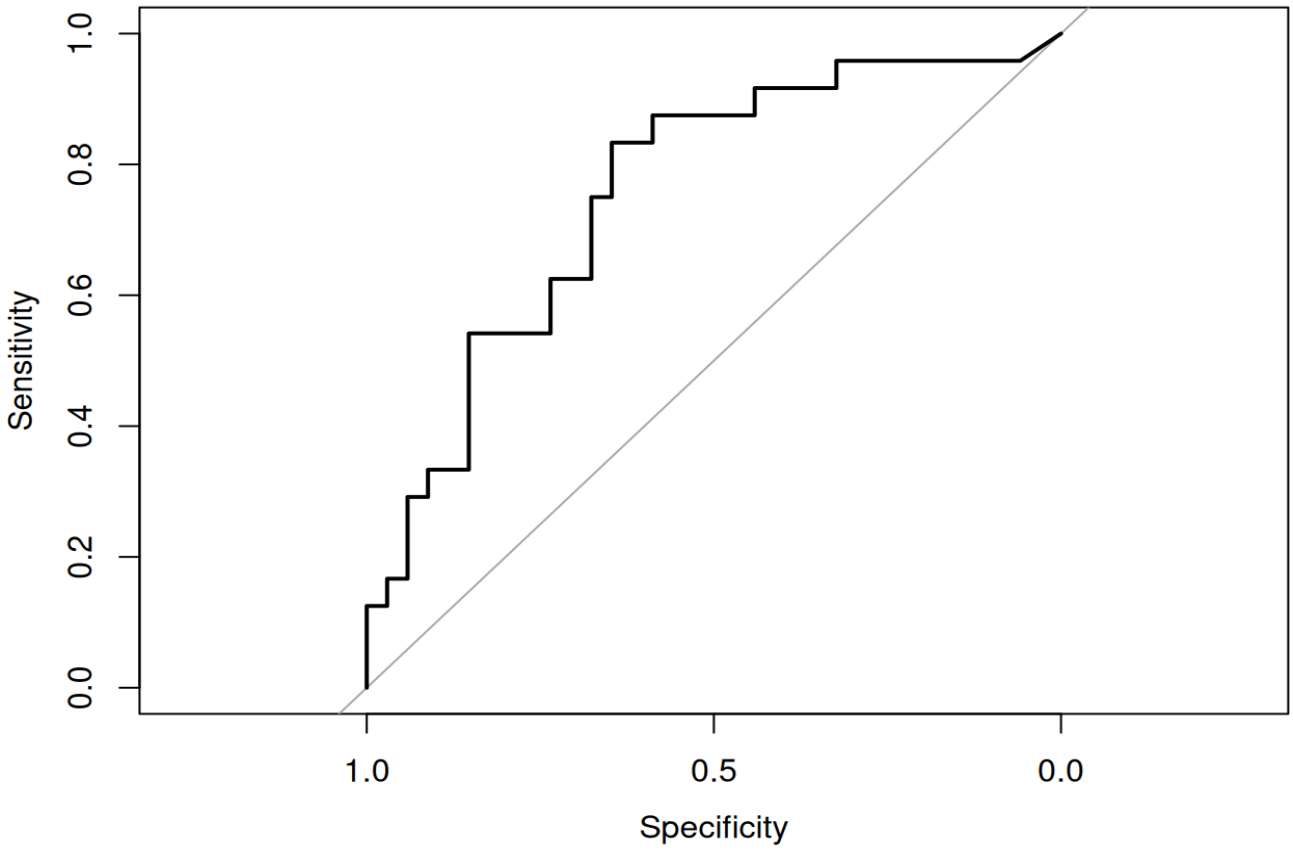
Volcano plots demonstrating differences in BF mRNA content between (2a) hCG negative and positive embryos and (2b) NPO and live-birth embryos using DESeq2. Differentially abundant genes were detected using DESeq2 and genes with a p-value <0.05 and a fold change of +2/-2 were considered statistically significant. Abbreviations: NPO=negative pregnancy outcome; BF=blastocoel fluid; hCG=human chorionic

Figure 3. Influence of maternal age on differentially abundant mRNA in BF



Volcano plots demonstrating differences in BF mRNA content between hCG negative and positive embryos from (3a) women older than 34 years and from (3b) women younger than 35 years. Differentially abundant genes were detected using DESeq2 and genes with a p-value <0.05 and a fold change of +2/-2 were considered statistically significant. Abbreviations: BF=blastocoel fluid; hCG=human chorionic gonadotropin.

Figure 4. Prediction model of pregnancy result based on mRNA in BF



ROC curve demonstrating the sensitivity and specificity of the most accurate ML model for prediction of pregnancy meaused as a positive or negatuve urine hCG 14 days after embryo transfer. The prediction model is based on machine learning using L2RLR with DAGs as input feature and PCA for pre-processing. Abbreviations: BF=blastocoel fluid; L2RLR=L2 Regularized Logistic Regression; DAG=differentially abudant genes; PCA=principal component analysis; hCG=human chorionic gonadotropin.