Amniotic fluid metabolic biomarkers of fetal physiology and pregnancy success

Victor A. Absalon-Medina[1,2], Rodrigo V. Sala[1], Daniela C. Pereira[3], Vanessa C. Fricke[1], Iebu Devkota[4], Zachary L. Bonomo[4], Dailin M. Fuego[4], Michael McDonald[5], José M. Sánchez[6], Maria B. Rabaglino[5,7], Antonios Matsakas[8], Anastasios Vourekas[9], Xing Fu[4], Rocio M. Rivera[10], Patrick Lonergan[5], Pablo J. Ross[3,11] & Constantine A. Simintiras[4,*].

[1] ST Genetics, Ohio Heifer Center, South Charleston, OH, 45368, USA
[2] Department of Animal Sciences, The Ohio State University, Columbus, OH, 43210, USA
[3] ST Genetics, Navasota, TX, 77868, USA
[4] School of Animal Sciences, Agricultural Center, Louisiana State University, Baton Rouge, LA, 70803, USA
[5] School of Agriculture and Food Science, University College Dublin, Belfield, Dublin, C04V1W8, Ireland
[6] Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, 28040, Spain
[7] Department of Population Health Sciences, School of Veterinary Medicine, Utrecht University, Utrecht, 3584, Netherlands
[8] Department of Life Sciences, Manchester Metropolitan University, Manchester, M1 5GD, United Kingdom.
[9] Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, 70803, USA
[10] Division of Animal Sciences, University of Missouri, Columbia, MO, 65211, USA
[11] Department of Animal Science, University of California, Davis, CA, 95616, USA
[*] Correspondence: csimintiras@agcenter.lsu.edu

ABSTRACT

Amniotic fluid (**AF**) profiling provides a minimally invasive window into early fetal physiology. We characterized the AF metabolome from first trimester (Day 68) Holstein dairy heifers (n=45), considering fetal sex, conception method [*in vitro* fertilization (**IVF**) *vs*. artificial insemination (**AI**)], and eventual pregnancy outcome as key variables. Multivariate statistics uncovered differentially abundant metabolites for each comparison – including markers that preceded spontaneous abortion – independently of recipient age, weight, gestation length, or fetal genetics. Thereafter, a machine learning algorithm using panels of six metabolites accurately predicted fetal sex (AUROC=0.76; *P*=0.023) and pregnancy viability (AUROC=0.81; *P*=0.018), while corroborating conception method (AUROC=0.91; *P*=0.001). External validation using AF (Day 42) from an independent cohort of beef heifers (n=22) reproduced the fetal sex classifier with similarly high sensitivity and specificity (AUROC=0.85, *P*=0.029). These findings reveal metabolic signatures that forecast fetal sex and pregnancy viability, while confirming distinct metabolic imprints of assisted-conception modalities. These data lay the groundwork for next-generation AF prenatal diagnostics in veterinary and human obstetrics.

KEYWORDS

Amniocentesis
Amniotic fluid
Bovine
Fetal metabolism
Fetal physiology
Machine learning
Metabolomics
Pregnancy diagnostics

INTRODUCTION

Amniocentesis has enabled fetal genetic diagnosis in cattle and humans for decades [1–4]. However, the broader diagnostic potential of amniotic-fluid (**AF**) composition remains largely unexplored. While AF is broadly isosmotic with fetal serum and undergoes gradual compositional changes across gestation [5], it remains the primary conduit for maternal-fetal biochemical exchange throughout gestation [6–9]. Accordingly, our overarching hypothesis was that comprehensive metabolomic profiling of this readily accessible fluid – obtainable through established, low-risk amniocentesis [10–12] – may provide new insights to refine prenatal monitoring in livestock and, potentially, human medicine.

The USA cattle production industry, valued at ~ $ 88 billion annually [13], faces reproductive efficiency challenges. Despite fertilization rates exceeding 80%, early pregnancy losses approach 45% in dairy operations [14]. This paradox directly undermines farm profitability, where reproductive performance is a core economic driver [15–17]. Ultrasound, the current pregnancy surveillance standard, reveals gross fetal anatomy but cannot detect sub-clinical metabolic disturbances and remains relatively operator dependent.

Meanwhile, human AF research is constrained by population heterogeneity and confounding variables, such as maternal age, weight, ethnicity, and conception method [18–26]. Rodent models offer experimental control but require pooling samples across multiple fetuses and litters, compromising individual-level resolution [27,28]. Rodent models are also highly inbred, poly-ovulatory, and exhibit much shorter gestation lengths, which complicate data extrapolation to bovine and human pregnancies [29].

In contrast, bovine pregnancies provide ample AF volumes for individual fetus-level analysis. Cattle also share key reproductive characteristics with humans, including mono-ovulation, comparable gestation length, and an estrous cycle more broadly analogous to the menstrual cycle [30–32]. Moreover, bovine embryonic epigenetic patterning more closely resembles human patterns than murine models [33,34] while experimental conditions, including genetics, nutrition, and environment can be effectively controlled.

We therefore applied ultra-high-throughput untargeted metabolomics, integrated with machine-learning analytics, to interrogate early-gestation bovine AF, addressing three specific objectives.

We first investigated whether AF contains metabolomic signatures of fetal sex – information valuable for livestock management [35]. Despite evidence that sex influences embryonic

metabolism [36–42] and epigenetic patterning [43–45], bovine AF sexual dimorphism has been examined in only one miRNA study [46].

Secondly, we examined metabolic differences between *in vitro* fertilization (**IVF**) and artificial insemination (**AI**) derived pregnancies. *In vitro* embryo production (**IVEP**) now dominates commercial cattle breeding [47], while human IVF exceeds 2.5 million cycles annually [48]. Understanding how assisted reproduction alters fetal metabolism [49–51], fetal epigenetics [52–57], and endometrial responses [58,59] could improve protocols for both species.

Thirdly, we tested whether AF metabolites can forecast pregnancy viability. Previous research identified four amino acids distinguishing viable from non-viable bovine pregnancies following cloned embryo transfer [60]. However, the chromosomal instability of clones [61] limits practical application. We therefore focused on spontaneous losses under standard commercial (IVF and AI) conditions.

Here we show that bovine AF (Day 68) harbors metabolomic signatures that can predict fetal sex and pregnancy outcome while corroborating conception method – independently of several maternal traits and fetal genetics. *Random Forest* models built on six metabolites each – including adenine, hypotaurine, methylguanosine, and phosphoserine, among others – achieved high *area under the receiver operating characteristic curve* (**AUROC**) values between 0.76-0.91. Furthermore, validation of the fetal sex classifier in an independent cohort confirmed model robustness. These findings highlight concise panels with potential near-future application in precision livestock management, where early and accurate prediction of calf sex and pregnancy trajectory could inform breeding decisions, optimize resource allocation, and reduce economic losses associated with undesired male calves or failed pregnancies. More broadly, this approach and these data provide a translational framework for developing next-generation prenatal diagnostics in human obstetrics.

MATERIALS AND METHODS

Overview

To systematically investigate AF metabolomic signatures, we established two experimental cohorts. Initial pregnancies were generated in Holstein heifers (n=20) at the ST Genetics Ohio Heifer Center (South Charleston, OH, USA) following estrous cycle synchronization and AI using conventional (n=13) or sex-sorted (n=7) semen (Cohort 1A). During pregnancy, ovum pickup (**OPU**) was performed on seven of these animals. Resulting oocytes were used for IVEP.

These embryos were individually transferred into a separate group of synchronized recipient Holstein heifers, generating a further 25 pregnancies (Cohort 1B). Amniocentesis was performed on all animals (n=45) on Day 68 of pregnancy. Cohort 1 animals were maintained by *ad libitum* access to a standard total mixed ration (TMR) corn silage-based diet. Cohort 1 samples were collected between 04.2020-12.2021 as approved by The Ohio State University Institutional Animal Care and Use Committee.

For independent validation, 22 embryos were transferred individually into estrous synchronized crossbred beef heifers (n=22) at the University College Dublin (**UCD**) Lyons Research Farm, Dublin, Ireland. AF was recovered on Day 42 (Cohort 2). Cohort 2 animals were maintained on a grass maize silage supplemented with a standard beef finishing concentrate. Cohort 2 samples were collected between 03.2022-08.2022 as approved by the UCD Animal Research Ethics Committee and licensed by the Health Products Regulatory Authority, Ireland, under Directive 2010/63/EU.

Experimental design

Cohort 1 AF samples were categorized according to multiple parameters to enable comprehensive analyses. This is summarized in **Figure 1** and described below.

Fetal sex. Male-carrying pregnancies (n=26) included fetuses derived by IVF (61.5%) and AI (38.5%), utilizing conventional (84.6%) and sex-sorted (15.4%) semen. Among these, 96.2% resulted in successful pregnancies, while 3.8% spontaneously aborted. Female-carrying pregnancies (n=19) consisted of IVF (47.4%) and AI (52.6%) derived fetuses, derived using conventional (84.2%) and sex-sorted (15.8%) semen. In this group, 73.7% carried to term, while 26.3% spontaneously aborted.

Conception method. IVF-derived pregnancies (n=25) comprised male (65%) and female (35%) fetuses, generated using conventional semen (100%) of which 88% resulted in successful pregnancy and 22% spontaneously aborted. AI-derived pregnancies (n=20) comprised male (50%) and female (50%) fetuses, generated using both conventional (65%) and sex-sorted (35%) semen. Of these, 85% were successful compared to 15% spontaneous abortions.

Pregnancy outcome. Among successful pregnancies (n=39), 64.1% were male and 35.9% were female. Of these, 56.4% were IVF-derived, and 43.6% were AI-derived. Additionally, 87.2% were produced using conventional semen, while 12.8% used sex-sorted semen. Spontaneous abortion (n=6) or successful pregnancy (n=39). The spontaneous abortion group (n=6) comprised 16.7% male and 83.3% female fetuses, with 50% derived from IVF and 50% from AI.

Of these, 66.7% were generated using conventional semen, and 33.3% used sex-sorted semen. Spontaneous abortions occurred at 208 ± 55.7 days (mean ± SD), ranging from Day 98-252.

Fetal genetics. Using semen from one sire throughout (controlling paternal effects) and collecting oocytes from seven heifers for IVEP allowed tracking genetic relationships. For example, oocytes from mothers producing Heifer 1 via AI also produced embryos transferred to recipients resulting in Heifers 2-7. Therefore, Heifers 1-7 (Group A) are full genetic siblings. This applies to Groups B-G, while Group H comprises paternal half-siblings (**Fig. 1D**).

Cohort 1A estrous cycle synchronization

Estrous cycles of 20 Holstein heifers were synchronized using a standard 5-day fixed-time artificial insemination (**FTAI**) protocol [62]. In brief, each heifer received a progesterone (**P4**)-controlled internal drug release (**CIDR**) device (1.38 g P4, Eazi-Breed, Zoetis, Florham Park, NJ) inserted intravaginally on a random day of their estrous cycle, plus intramuscular administration of gonadotropin releasing hormone (**GnRH**, 100µg gonadorelin acetate, Parnell, Overland Park, KS), designated as Day -8. On Day -3, the CIDR was removed, and heifers were administered prostaglandin F2α (**PGF2α**) intramuscularly (500 µg cloprostenol sodium, Parnell, Overland Park, KS). A second, identical PGF2α injection followed 24 hours later. The day of observed estrus was marked as Day 0 at which time **GnRH** was administered intramuscularly to induce ovulation (**Supplementary Figure 1A**).

Cohort 1A artificial insemination

Thirteen of these synchronized heifers, selected at random, were artificially inseminated on Day 0 using conventional bull semen from the same sire (ST Genetics, Navasota, TX). More specifically, semen was thawed by immersion in 35.5 °C water for 45 seconds before deposition into the uterine cavity, guided by transrectal palpation. The remaining seven synchronized heifers were identically artificially inseminated on Day 0 using sex-sorted semen from the same sire as previously (SexedULTRA 4M™, ST Genetics, Navasota, TX).

Cohort 1A amniocentesis

Pregnancies were confirmed on gestational Day 60 by transrectal ultrasonography using a 5-9 MHz linear transducer coupled to an Ibex EVO II display (E.I. Medical Imaging, Loveland, CO). Amniocentesis was then performed on Day 68, following a previously described procedure [63]. In brief, heifers were restrained in a squeeze chute, and gentle massage of the ventral vulvar area stimulated urination. An epidural block was administered by injecting 5 ml of 2 % lidocaine (Aspen Veterinary Resources, Liberty, MO) into the inter-coccygeal space between the first and

142 second vertebrae. Additionally, 10 mg xylazine (Rompun, Shawnee Mission, KS) was
143 administered intravenously as a sedative to further minimize stress.

144 A vaginal lavage was then performed by intravaginal infusion of 60 ml sterile 0.9 % sodium
145 chloride solution, while rectal contents were emptied to allow manipulation of the broad
146 ligament. After aseptic preparation, a 5-10 MHz convex transducer coupled to an Ibex EVO II
147 display (E.I. Medical Imaging, Loveland, CO) was inserted into the vaginal canal up to the fornix.
148 Using transrectal manipulation of the broad ligament, the amniotic space was positioned against
149 the vaginal wall to enable safe insertion of a 20 G x 2" needle (WTA, College Station, TX) to
150 minimize risk of injury to the umbilical cord, placentomes, or fetus.

151 The amniocentesis needle was connected to 1.4 m tubing (WTA, College Station, TX) and a 3-
152 way stopcock (MILA International, Florence, KY) with two syringes attached to the remaining
153 ports. Upon entering the amniotic cavity, 5 ml AF was drawn into a 20 ml luer-lock syringe to
154 prime the line. The port was then switched to collect a final AF volume of approximately 40 ml
155 within a 50 ml luer-lock syringe (Air-Tite Products Co., Inc., Virginia Beach, VA). AF samples
156 were immediately aliquoted, snap frozen in liquid nitrogen [$\mathbf{N_2}(\boldsymbol{l})$], and stored in $N_2(l)$ until
157 transport for analysis.

158 Following amniocentesis, pain management was provided by intravenous administration of
159 flunixin meglumine (Vetameg, Aspen Veterinary Resources, Liberty, MO) at 50 mg·50kg^{-1} body
160 weight, and oral administration of meloxicam (Unichem Pharmaceuticals Inc., East Brunswick,
161 NJ) at 50 mg·50kg^{-1} body weight. All amniocentesis procedures were performed by one of two
162 experienced technicians.

Cohort 1A ovum pickup

163 During pregnancy, OPU was performed on seven heifers using a standard protocol [64,65].
164 Specifically, donors were restrained in a squeeze chute, and caudal epidural anesthesia was
165 administered as described above. The perineal area was then cleaned and disinfected using 70
166 % isopropyl alcohol. Oocyte retrieval was ultrasound-guided using a 5-9 MHz linear transducer,
167 coupled to an Ibex EVO II display (E.I. Medical Imaging, Loveland, CO), inserted into the
168 vaginal fornix. Follicles were punctured with an 18 G x 5.5 cm needle (WTA, College Station,
169 TX) attached to a metal guide connected via plastic tubing to a 50 ml conical tube linked to a
170 vacuum pump (Cook Medical, Bloomington, IN), maintaining a constant flow rate of 12 ml·min^{-1}.

171 Plastic tubing was flushed with pre-warmed (38.5 °C) Dulbecco's phosphate buffer solution
172 (**DPBS**) supplemented with 0.4% bovine serum albumin (**BSA**), 25 mg·l^{-1} kanamycin sulphate

173 and 5 IU·ml^{-1} sodium heparin at 36 ± 1 °C. Then, follicular aspirates from each follicle over 3
174 mm in diameter were filtered using a 75 µm filter (oocyte aspiration dish with 3 mm grid,
175 Professional Embryo Transfer Supply, Canton, TX) and washed with pre-warmed oocyte
176 collection medium (Boviteq, Madison, WI). The contents of the filter were transferred to a square
177 grid dish to locate and harvest cumulus-oocyte complexes (**COC**) under a stereomicroscope.
178 OPU was performed by one of two experienced technicians.

Cohort 1B in vitro embryo production

179 *In vitro* embryo production was performed at the ST Genetics Texas laboratory following
180 proprietary procedures. COC from each donor were matured in-transit over 24 h at 38.5 °C
181 within maturation medium (ST Genetics, TX). Matured COC were then transferred into a pre-
182 equilibrated 60 µl drop of IVF medium (ST Genetics, TX) covered with mineral oil. Frozen-
183 thawed sperm was purified using a double-density gradient approach (Nidacon International AB,
184 Mölndal, Sweden) as previously described [66]. A final concentration of 10^6 sperm·ml^{-1} was
185 achieved and fertilization took place for 8 h at 38.5 °C under 5 % CO_2 in air. Following
186 fertilization, cumulus cells were removed, and embryos were cultured in a benchtop incubator
187 (WTA, College Station, TX) at 38.5 °C under 5 % O_2, 5 % CO_2, and balanced N_2 as premixed
188 gas (Airgas, Dallas, TX). Cleavage rates were recorded 3 days after fertilization and blastocyst
189 rates recorded at 7 days post-fertilization. Embryo stage and quality were morphologically
190 determined according to International Embryo Transfer Society (**IETS**) guidelines.

Cohort 1B embryo transfer

191 Recipients for embryo transfer were synchronized using a standard 5-day fixed-time embryo
192 transfer (**FTET**) protocol [62]. The heifers received a CIDR inserted intravaginally on a random
193 day of their estrous cycle designated as Day -8. On Day -3, the CIDR was removed, and the
194 heifers were administered an intramuscular injection of PGF2α. After 72 hours GnRH was
195 administered intramuscularly to induce ovulation at which time marked as Day 0
196 (**Supplementary Figure 1B**). The presence of a *corpus luteum* (**CL**) on Day 5 was confirmed
197 by transrectal ultrasonography. On Day 7 of the estrous cycle, Day 7 embryos were loaded
198 individually into 0.25 ml French straws with holding medium (ST Genetics, TX). Loaded straws
199 were placed in a portable incubator (Micro Q Technologies, Scottsdale, AZ) at 38.5 °C and
200 transported to the farm. Prior to embryo transfer, regional anesthesia was administered through
201 a caudal epidural injection as aforementioned. Each recipient received a single blastocyst,
202 which was transferred into the uterine horn ipsilateral to the CL. Pregnancy was confirmed by

203 transrectal ultrasonography on Day 60 post-transfer and amniocentesis was performed on Day
204 68 as described above.

Cohort 2 estrous synchronization

205 Genetically unrelated crossbred beef heifers (n=22), primarily Limousin, Charolais, or Aberdeen
206 Angus crosses were synchronized using an analogous and established [67–70] protocol
207 (**Supplementary Figure 1C**). In brief, heifers received a P4 releasing intravaginal device (**PRID**
208 **E**) (1.55 g P4, Ceva Santé Animale, Libourne, France) on a random day of their estrous cycle,
209 designated as Day -11, concomitantly with GnRH analogue (Ovarelin, 100 mg gonadorelin;
210 Ceva Santé Animale) intramuscular administration. After seven days, PGF2α (Enzaprost, 5 mL
211 equivalent to 25 mg dinoprost; Ceva Santé Animale) was administered, before PRID removal
212 the following day.

Cohort 2 in vitro embryo production

213 Blastocysts were produced *in vitro* using an analogous and established protocol [71]. Briefly,
214 immature COC were collected by aspirating follicles from the ovaries of cattle slaughtered at a
215 local abattoir (Kildare Chilling Company, Kildare, Ireland). COC were pooled, washed in PBS,
216 and matured for 24 h in groups of 50 in 500 µl of TCM-199 (Sigma Aldrich, Arklow, Ireland),
217 supplemented with 10 % fetal calf serum and 10 $ng \cdot ml^{-1}$ epidermal growth factor (Sigma
218 Aldrich). Maturation took place at 39 °C in a humidified environment with 5 % CO_2 in air.

219 Mature COC were inseminated with frozen-thawed sperm (National Cattle Breeding Centre,
220 Kildare, Ireland) at a concentration of 10^6 $sperm \cdot ml^{-1}$. After 20 h of co-incubation at 39 °C under
221 5 % CO_2 in air, presumptive zygotes were denuded by vortex and cultured in 25 µl droplets of
222 IVC medium (Stroebech), supplemented with 3 $mg \cdot ml^{-1}$ bovine serum albumin (Sigma Aldrich)
223 at 39°C in a humidified atmosphere with 5 % CO_2 and 5 % O_2 under mineral oil. Embryos were
224 cultured at a ratio of 1 $embryo \cdot \mu l^{-1}$. Grade 1 blastocysts for transfer were collected on Day 7 and
225 loaded into straws with embryo holding medium (IMV Technologies, L'Aigle, France).

Cohort 2 embryo transfer

226 Heifers were monitored for signs of estrus five times daily, starting 30 h after PRID withdrawal.
227 All heifers observed standing estrus and thus received a single Day 7 *in vitro*-produced
228 blastocyst on Day 7 or 9 of the estrous cycle, with Day 0 being considered the day of expected
229 ovulation (approximately 28 h after estrus onset). Embryo transfer was performed as described
230 above. Pregnancies were confirmed by transrectal ultrasonography on Day 28 of gestation, and
231 all transfers were conducted by one of two experienced technicians.

Cohort 2 amniotic fluid recovery

232 AF was collected as previously described [72]. In brief, pregnant heifers were slaughtered on
233 Day 42 of gestation in a commercial European Union licensed abattoir. The reproductive tract
234 was recovered and kept on ice until processing for sample collection, within 30 min of slaughter.
235 The pregnant uterine horn was opened along the major curvature to retrieve fetal membranes.
236 For AF collection, a 30 G needle connected to a 1 ml syringe was used to pierce the amnion
237 and aspirate the fluid. AF was placed into RNase/DNase-free tubes (Thermo Fisher Scientific,
238 Waltham, USA), centrifuged at 16,000 *x g* for 10 min at 4 °C, and the supernatant placed into
239 new RNase/DNase-free tubes, snap-frozen in liquid nitrogen, and stored at -80 °C until analysis

Dependent experimental variable metrics

240 Additional animal details and dependent experimental variable raw metrics are provided in
241 **Supplementary Figure 1D** (Cohort 1) and **Supplementary Figure 1E** (Cohort 2).

Mass spectrometry

242 AF samples were first thawed on ice for 60 min, vortexed, and centrifuged briefly to remove
243 bubbles. After addition of 180 µl 80% methanol per 20 µl of each sample, samples were
244 incubated at 4°C for 1 h and then centrifuged at 3,220 *x g* for 15 min at room temperature. The
245 resulting supernatants were stored at -80°C until analysis. Therefore, these data represent the
246 methanol-extractable portion of the amniotic fluid metabolome.

247 Metabolic profiling was conducted at General Metabolics Inc. (Boston, MA) using flow-injection
248 mass spectrometry (**FI-MS**) on an Agilent 6550 quadrupole time-of-flight (**Q-TOF**) system [73],
249 similarly to Chen *et al*. [74]. In brief, the equipment was configured to scan in full MS mode at
250 1.4 Hz, operating in negative ionization with 4 GHz high resolution mode, across a mass range
251 of 50 to 1,000 m/z. The solvent, 60% isopropanol, was supplemented with 1 mM ammonium
252 fluoride (NH_4F) at pH 9.0, 10 nM hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazene, and 80
253 nM taurocholic acid, for mass calibration.

254 Samples (100 µl each) were injected in randomized order and data were acquired in profile
255 mode. Data were centroided before analysis using MATLAB (MathWorks). Missing values were
256 imputed using recursive analysis, and consensus centroids were identified across all samples.
257 Ions were annotated based on accurate mass and isotopic patterns using the HMDB database
258 (version 4.0) [75]. It is worth noting that due to the inherently weak chromatographic separation
259 often associated with global metabolomic profiling [76], compounds with identical molecular

formulae could not be distinguished. Therefore, annotation confidence was level 4, though, in practice, it is generally higher for common metabolites [73].

A pooled aliquot of all experimental samples, serving as a technical replicate control, was run in between the experimental samples at defined intervals. Based on these, the mean technical (instrument) standard error was calculated at 2.8 %.

Metabolomic analyses

Single-factor analyses were conducted using MetaboAnalyst 6.0 [77]. Initially, raw peak intensities were filtered by interquartile range to account for variance, following standard recommendations for untargeted metabolomics datasets [78]. Based on our mean technical standard error, a threshold of 5% was applied, leading to the exclusion of 68 metabolites. Data were then normalized to the median, log-transformed (base 10), and auto-scaled (mean-centered and divided by each variable standard deviation).

Thereafter, volcano plots were generated by unpaired t-test, with thresholds set at a $P \leq 0.05$, a fold change of 1.0 (*i.e*. no change), and assuming equal group variance. Principal component analysis (**PCA**) plots with 95% confidence intervals were created using permutational multivariate analysis of variance (PERMANOVA), with distributions based on Euclidean distance from the first two principal components.

Sparse partial least squares discriminant analysis (**sPLS-DA**) was performed with 5 components and 10 variables per component. Model performance was evaluated using 5-fold cross-validation with an increasing number of components and a fixed 10 variables per component. Hierarchical clustering dendrograms were generated using Euclidean distances and the Ward method. Heatmaps were produced from normalized data, standardized by auto-scaled metabolite features, with Euclidean distance and Ward clustering applied.

For metadata analyses, peak intensities were filtered, normalized, transformed, and scaled as described above. A metadata heatmap was generated using Euclidean distance and Ward clustering for both metabolites and metadata variables. Correlation coefficients were calculated using the Pearson R correlation measure. Linear models with covariate adjustment ($P \leq 0.05$) were applied using the *limma* linear regression approach, as previously described [76,79].

For categorical enrichment analyses, compound names were first standardized against the HMDB, PubChem, and KEGG databases. Unstandardized compound names were excluded. Peak intensities were then normalized, transformed, and scaled as in prior steps. Enrichment testing was conducted based on the *global test* [80] against the RaMP-DB metabolite set library,

291 which integrates 3,694 features from KEGG (via HMDB), Reactome, and WikiPathways
292 databases. Only metabolite sets with at least two entries were included. Enrichment values
293 were calculated as the ratio of observed *vs*. expected metabolites within each pathway [81].

294 In contrast, pathway topology analyses were conducted by first standardizing compound names
295 against the HMDB, PubChem, and KEGG databases, before peak intensity normalization,
296 transformation, and scaling – all as above. The pathway analysis focused on significant
297 ($P \leq 0.05$) metabolites rather than pre-selected ones, with enrichment performed using the *global*
298 *test* as above. Topology was assessed using relative-betweenness centrality, and the reference
299 metabolome included all compounds from the KEGG *Bos taurus* library. Scatter plots were
300 generated to display all matched pathways, with *P*-values from the pathway enrichment analysis
301 plotted against pathway impact values from the topology analysis.

Machine-learning based biomarker identification

302 Biomarker analyses were performed using MetaboAnalyst 6.0, leveraging the receiver operating
303 characteristic (**ROC**) curve-based model evaluation function. Initially, raw peak intensities were
304 filtered, normalized, transformed, and scaled as described above. Metabolites were then
305 manually selected for ROC analysis, which was conducted using the *Random Forests*
306 multivariate algorithm. Specifically, 100 cross-validations were performed, with results averaged
307 to generate ROC curves with 95% confidence intervals and predictive accuracy values.
308 Empirical *P*-values were calculated from 1,000 AUROC permutations.

RESULTS

Metadata summary

309 Ultra-high-throughput untargeted metabolomic profiling of AF collected from Holstein heifers
310 (n=45) on Day 68 of pregnancy (Cohort 1) identified 1,358 metabolites, with 1,335 annotated
311 (**Supplementary Table 1**). Following variance filtration, 68 metabolites (5%) were excluded,
312 leaving 1,290 metabolites for analysis. The mean (± SD) age (591.5 ± 73.7 days) and weight
313 (502.4 ± 74.5 kg) of recipient heifers at the time of amniocentesis were similar, and among
314 successful pregnancies, gestation length (274.5 ± 4.9 days) and calf birth weight (41.9 ± 7.7 kg)
315 were comparable (**Supplementary Figure 1D**).

Amniotic fluid composition is sexually dimorphic

316 We first compared the AF metabolome from male *vs*. female fetuses. Initial principal component
317 (**PCA**; **Supplementary Figure 2A**) and hierarchical clustering (**Supplementary Figure 2B**)

318 analyses revealed high AF composition overlap between both groups. As such, the overall AF
319 metabolic landscapes between male *vs*. female fetuses are very similar. However, subsequent
320 sparse partial least squares discriminant analysis (**sPLS-DA**) – a supervised machine learning
321 algorithm effective in separating non-linear clustered signals [82–84] – segregated the AF
322 metabolic signatures of the two groups (**Fig. 2A**), albeit with a high mean cross-validation (**CV**)
323 error rate of 42.2 % (**Supplementary Figure 2C**).

324 AF metabolome correlations between fetal sex and (*a*) discrete covariates [pregnancy outcome
325 (**Fig. 2B**), conception method (**Fig. 2C**), semen source (**Fig. 2D**), fetal genetics (**Fig. 2E**)], and
326 (*b*) continuous variable metadata [recipient weight (**Fig. 2F**), recipient age (**Fig. 2G**), gestation
327 length (**Fig. 2H**), offspring birthweight (**Fig. 2I**)] variables were low, with a mean (± SD)
328 correlation coefficient (*R*) of 0.12 ± 0.2. This demonstrates no significant association between
329 these parameters and fetal sex in terms of AF composition.

330 Before covariate adjustment, 15 metabolite (1.2 % of total) relative concentrations differed
331 ($P \leq 0.05$) between AF from male and female fetuses (**Fig. 2J-K**). Qualitative enrichment analysis
332 highlighted tRNA-derived modified nucleoside, estrone, and estrogen metabolism as over-
333 represented pathways (**Fig. 2L**). A semi-quantitative targeted pathway analysis further revealed
334 metabolites corresponding to α-linolenic acid and unsaturated fatty acid metabolism as
335 differentially abundant in the same comparison (**Supplementary Figure 2D**).

336 Subsequent linear covariate adjustment metabolomic analysis, to increase precision and reduce
337 bias [85], isolated 13 differentially abundant ($P \leq 0.05$) AF metabolites (1 % of total) by fetal sex –
338 independently of fetal conception method, semen type, pregnancy outcome, and fetal genetics
339 (**Fig. 2M**). These include hexonic acid, isoeugenol phenylacetate, and methylguanosine. **Figure**
340 **2N** further highlights the lack of association between these metabolite concentrations and
341 recipient age, weight, and gestation length. Therefore, despite high overlap between AF
342 metabolomes from male and female fetuses, there are select differences, underscoring sexually
343 dimorphic fetal metabolism.

In vitro fertilization alters amniotic fluid composition

344 We next compared the AF metabolome from fetuses derived using IVF *vs*. AI. Initial PCA
345 (**Supplementary Figure 3A**) and hierarchical clustering (**Supplementary Figure 3B**) revealed
346 similarly high AF composition overlap between both groups, which could be overcome using
347 sPLS-DA (**Fig. 3A**). The associated mean sPLS-DA CV error was 23.6 % (**Supplementary**
348 **Figure 3C**). Therefore, consistent with fetal sex, the overall AF metabolic profiles of IVF- and AI-
349 derived fetuses are similar, but not identical.

350 AF metabolome correlations between fetal conception method and (a) discrete covariate
351 [pregnancy outcome (**Fig. 3B**), fetal sex (**Fig. 3C**), semen source (**Fig. 3D**), fetal genetics (**Fig.**
352 **3E**)], and (b) continuous variable metadata [recipient weight (**Fig. 3F**), recipient age (**Fig. 3G**),
353 gestation length (**Fig. 3H**), offspring birthweight (**Fig. 3I**)] variables were low. More specifically, a
354 mean (± SD) correlation coefficient (*R*) of 0.05 ± 0.40 indicates no significant association
355 between these parameters and conception method in terms of AF composition.

356 Prior to covariate adjustment, 49 metabolite (3.8 % of total) relative concentrations differed
357 ($P \leq 0.05$) between AF from AI- and IVF-derived fetuses (**Fig. 3J-K**). Qualitative enrichment
358 analysis highlighted three pathways related to pyrimidine metabolism, three related to
359 inflammation, and three related to ceramide signaling as over-represented (**Fig. 3L**). Semi-
360 quantitative targeted pathway analysis further confirmed pyrimidine and sphingolipid (including
361 ceramide) metabolism impact (**Supplementary Figure 3D**).

362 After similar covariate adjustment, 12 AF metabolites (0.9 % of total) were differentially
363 abundant ($P \leq 0.05$) based on fetal conception method – independently of fetal sex, semen type,
364 pregnancy outcome, and fetal genetics (**Fig. 3M**). These include thymine, C16:3, and
365 oxocortisol. **Figure 3N** further highlights the lack of association between these metabolite
366 relative concentrations and continuous metadata. Thus, much like with fetal sex, there is
367 significant overlap in the AF metabolomes of AI and IVF-derived fetuses; however, specific
368 differences highlight a metabolic impact of IVF.

Pregnancy outcome is reflected in the amniotic fluid metabolome

369 Next, we retrospectively analyzed the AF metabolome from successful and spontaneously
370 aborted pregnancies. Like previously, PCA (**Supplementary Figure 4A**) and hierarchical
371 clustering (**Supplementary Figure 4B**) showed high AF composition overlap between both
372 groups, although sPLS-DA was able to differentiate the AF metabolomes from pregnancies of
373 divergent viability (**Fig. 4A**) with a lower mean CV error of 14.7 % (**Supplementary Figure 4C**).
374 Therefore, while the overall AF metabolic profiles of pregnancies of divergent viability are very
375 similar, select differences are apparent.

376 AF metabolome correlations between pregnancy outcome and (a) discrete covariate [fetal sex
377 (**Fig. 4B**), conception method (**Fig. 4C**), semen source (**Fig. 4D**), fetal genetics (**Fig. 4E**)], and
378 (b) continuous variable metadata [recipient weight (**Fig. 4F**), recipient age (**Fig. 4G**), gestation
379 length (**Fig. 4H**), offspring birthweight (**Fig. 4I**)] variables were similarly low, with a mean (± SD)
380 correlation coefficient (*R*) of 0.17 ± 0.32. This is unsurprising given that gestation lengths and
381 birth weights could only be determined from successful pregnancies. Incidentally, the strongest

correlation observed among successful pregnancies was between offspring birth weight and gestation length (*R*=0.716; **Supplementary Figure 5A**), as expected.

Before covariate adjustment, 35 metabolite (2.7 % of total) relative concentrations differed (*P*≤0.05) between AF from successful *vs*. nonviable pregnancies (**Fig. 4J-K**). Qualitative enrichment analysis highlighted three pathways related to purine metabolism, two related to defective solute carriers, and two related to molybdenum cofactor imbalances as over-represented (**Fig. 4D**). This was largely reflected in a semi-quantitative targeted pathway analysis, as purine metabolism – to which molybdenum cofactor dependent enzymes are central [86] – ranked highly (**Supplementary Figure 4D**).

After covariate adjustment, 28 metabolites (2.2 % of total) were differentially abundant (*P*≤0.05) based on pregnancy viability, independently of discrete covariates (**Fig. 4M**). These include pentosidine, allantoin, and hydroxyphenylacetothiohydroximate (C17239). **Figure 4N** further highlights the lack of association between these metabolite relative concentrations and continuous metadata variables.

Thus, similarly to fetal sex and conception method, there is significant overlap in the AF metabolomes of fetuses of divergent competence; however, specific differences suggest these metabolites may be promising biomarkers of pregnancy viability. Furthermore, extrapolated metabolic pathways of likely importance to fetal survival are presented in **Figure 5**. These were manually constructed and primarily revolve around the tricarboxylic acid, urea, methionine, and purine nucleotide cycle interconversions.

Machine-learning based biomarker identification

We next used the receiver operating characteristics (**ROC**) curve with the *Random Forest* machine learning algorithm approach to highlight metabolite relative concentrations that may confirm fetal conception method and predict fetal sex and pregnancy outcome. Biomarker area under the ROC curve (**AUROC**) values can be generally classified as excellent (0.9-1.0), good (0.8-0.9), fair (0.7-0.8), poor (0.6-0.7), or fail (0.5-0.6) [87]. We also performed a cross-validation (**CV**) analysis to evaluate the predictive accuracy of each model.

The relative concentrations of six AF metabolites (**Fig. 6Ai**) could predict fetal sex with an AUROC of 0.76 (*P*=0.023) (**Fig. 6Aii**) and CV predictive accuracy of 0.68 (**Fig. 6Aiii**). Regarding conception method, the relative concentrations of six different AF metabolites (**Fig. 6Bi**) could confirm whether a fetus was derived by AI or IVF with an AUROC of 0.91 (*P*=0.001) (**Fig. 6Bii**) and CV of 0.83 (**Fig. 6Biii**). Finally, the relative concentrations of another six AF metabolites

413 (**Fig. 6Ci**) could predict whether a pregnancy would be successful or not with an AUROC of
414 0.81 (*P*=0.018) (**Fig. 6Cii**) and CV of 0.70 (**Fig. 6Ciii**).

Biomarker model confirmation

415 To further test these data, AF was collected from an independent group of crossbred beef
416 heifers (n=22) on Day 42 of pregnancy, carrying male or female fetuses (Cohort 2). Mean (±
417 SD) age (757.0 ± 58.6 days) and weight (586.1 ± 27.0 kg) of recipient heifers at the time of
418 amniocentesis were similar. However, mean Cohort 2 recipient ages (**Supplementary Figure**
419 **5B**) and weights (**Supplementary Figure 5C**) differed to Cohort 1 (*P*≤0.0001). Cohort 2
420 pregnancies were established by embryo transfer following IVF using conventional semen from
421 a different single sire to Cohort 1, producing 17 male and 5 female fetuses. AF from Cohort 2
422 was identically subjected to high-throughput untargeted metabolomics. Comparative analyses of
423 AF metabolomic profiles from both cohorts revealed greater variation between cohorts than
424 between fetal sexes, as indicated by hierarchical clustering (**Fig. 6D**) and sPLS-DA (**Fig. 6E**).
425 Despite this cohort-specific variation, the same six AF metabolites used to predict fetal sex in
426 Cohort 1 (**Fig. 6Fi**) achieved similarly high predictive performance in Cohort 2, with an AUROC
427 of 0.85 (P = 0.029) (**Fig. 6Fii**) and a cross-validated accuracy of 0.78 (**Fig. 6Fiii**).

DISCUSSION

428 These data provide compelling evidence that, while the overall AF metabolome is highly similar
429 across different conditions, specific metabolomic changes are associated with fetal sex,
430 conception method, and pregnancy outcome. These findings deepen our understanding of fetal
431 development and highlight potential biomarkers that could be used in reproductive management
432 and prenatal diagnostics, particularly for predicting mammalian fetal sex and pregnancy viability.

Amniotic fluid composition is sexually dimorphic

433 Sex-specific differences during development are well-established. However, a better
434 comprehension of the sexually dimorphic aspects of pregnancy is crucial for advancing
435 individualized prenatal care and elucidating sex-specific health trajectories. Here we found that
436 the *tRNA derived modified nucleosides* pathway was most differentially enriched in AF by fetal
437 sex, driven largely by elevated methylguanosine in female *vs*. male pregnancies. Methylation of
438 guanosine at the N^7 position [N^7-methylguanosine (**m^7G**)] at nucleotide 46 of tRNA (**m^7G-46**) is
439 one of the most prevalent and conserved tRNA modifications [88,89]. It plays a vital role in
440 regulating steady-state tRNA levels, which affect cell growth and behavior [90,91], including in

441 mammalian stem cells [92,93]. This methylation is catalyzed by the METTL1-WDR4 complex
442 [94]. Intriguingly, cell-free METTL1 mRNA transcript levels are reduced in AF from pregnancies
443 with Turner syndrome (gonadal dysgenesis) offspring [95], suggestive of a causal link between
444 tRNA guanosine methylation and sex-specific fetal development regulation.

445 The tRNA m^7G-46 modification is also implicated in autophagy regulation [96]. Sexually
446 dimorphic placental autophagy has been observed in response to stress [97,98]. Moreover,
447 miscarriage, which occurs more frequently in pregnancies with male fetuses [99–101], is often
448 closely associated with placental autophagy [102]. It is therefore plausible that the sexual
449 dimorphism in AF methylguanosine levels observed here arises from differential placental
450 regulation of m^7G-46 tRNA. This is consistent with the facts that, during the first-trimester,
451 human placental DNA methylation is sexually dimorphic [103], and AF composition primarily
452 reflects placental transudate from maternal circulation [104].

453 Given the numerous parallels between pregnancy and cancer [105], it is intriguing that m^7G is
454 also secreted by malignant cancer cells [106]. Cellular mechanisms influenced by m^7G-46
455 tRNA, common to both pregnancy and cancer, include immune evasion, proliferation, and
456 migration [107]. As such, sexually dimorphic methylguanosine levels in AF may partially reflect,
457 or contribute to, the phenomenon of accelerated male *vs*. female fetal development [108].
458 Further linking tRNA modifications to sexually dimorphic fetal development is the identification of
459 queuosine, a modified nucleoside found at nucleotides 34-37 (the 'wobble' positions [109]) of
460 specific tRNAs, in bovine AF [110], albeit not in this study. Queuosine-modified tRNAs have
461 been shown to promote sex-dependent learning and memory formation in mice [111], and
462 queuosine-modified tRNA glycosylation is required for post-embryonic growth in zebrafish [112].
463 Together, these findings implicate epitranscriptomic modifications in fetal development sexual
464 dimorphism. However, whether AF composition is a cause or effect of these modifications
465 remains an open question.

In vitro fertilization alters amniotic fluid composition

466 Elucidating the molecular determinants of altered development in a subset of offspring
467 conceived via IVF is of significant public health importance, particularly given the growing use of
468 assisted reproductive technologies [113]. More specifically, understanding these biomarkers
469 and mechanisms is crucial for enhancing clinical practices and reducing the risk of adverse
470 outcomes for offspring. After controlling for discrete confounding variables, we found 12
471 differentially abundant AF metabolites based on fetal conception method. The most significant
472 was thymine, elevated in the AF of AI *vs*. IVF pregnancies. Thymine derivatives are present in

urine as a byproduct of DNA damage repair (**DDR**) mechanism activity [114–116]. Thus, it is tempting to suggest that DDR is more efficient in AI *vs*. IVF pregnancies. Supporting this is a recent finding that plastics used in IVF alter placental gene expression in mice, with DNA repair being a main enriched gene set [117].

The elevated arginyl-glutamic acid (**Arg-Gln**) dipeptide in AI pregnancies further supports the hypothesis that natural conception provides metabolic advantages. Arg-Gln demonstrates protective effects on developing organs [118,119] and positively correlates with birth weight in women [120]. Given the economic importance of calf birth weight in cattle production, understanding these metabolomic differences could inform selection strategies for embryo transfer recipients and optimize protocols to maximize offspring viability.

Dipeptides are generally produced through aminopeptidase-catalyzed protein cleavage [121], with aminopeptidase M known to be present in human AF [122]. This raises the intriguing question of whether Arg-Gln is synthesized within AF itself. This would suggest that AF possesses a degree of metabolic semi-autonomy and would parallel the metabolic capabilities of uterine fluid, which has been shown to support specific biochemical processes independently [123]. If confirmed, the ability of AF to semi-autonomously metabolize metabolites such as Arg-Gln would suggest a more active role in fetal development and nutrient processing than previously understood, influencing fetal growth and development in novel ways.

Pregnancy outcome is reflected in the amniotic fluid metabolome

During this study, a subset of pregnancies from which AF was collected resulted in spontaneous abortion, providing an opportunity to investigate biomarkers related to pregnancy outcome. Identifying these markers is critical for the early detection of potential pregnancy complications and developing tailored interventions to improve pregnancy success in both women and cattle. This is particularly significant given that over 3 million stillbirths occur globally each year [124] and the economic significance of livestock reproductive management to agriculture.

We found that the *molybdenum cofactor biosynthesis* pathway was most differentially enriched in AF when stratified by subsequent pregnancy outcome. This is primarily attributable to reduced xanthine and elevated urate levels (amniotic hyperuricemia) in spontaneously aborted *vs*. successful pregnancies. Xanthine oxidase (**XO**), one of only four molybdenum cofactor enzymes in higher-order mammals [125], primarily catalyzes two reactions: the conversion of hypoxanthine to xanthine and the oxidation of xanthine to urate [126]. XO is derived from xanthine dehydrogenase under hypoxic conditions through phosphorylation [127]. Thus, it is unsurprising that intrapartum hypoxia – a condition associated with pregnancy loss [128] – is

linked to increased fetal XO activity [129] and, by extrapolation, amniotic hyperuricemia. This is supported by the finding that maternal administration of allopurinol, a XO inhibitor that transverses the placenta in several species [130–133], generally yields positive outcomes in pregnancies complicated by intrapartum hypoxia [133–138]. While intrapartum hypoxia occurs in 0.1 to 1.2 % of human pregnancies [139], our findings indicate that it may be more prevalent in cattle. This possibility is further supported by human data showing that second-trimester amniotic hyperuricemia predicts infant birth weight in normotensive women, independently of many systemic maternal factors such as insulin resistance [140].

Although the specific causes of spontaneous abortion in this study are unknown, our data suggest that disrupted AF urate homeostasis is associated with adverse pregnancy outcomes. These findings also indicate that aspects of pregnancy loss may be conserved between women and cattle. Moreover, several differentially abundant AF metabolites identified in this study are metabolically related (**Fig. 5**), suggesting shared pathways that may be linked to pregnancy failure. However, at this point this is quite speculative; thus, further research is needed to confirm these potential causal associations. It is also worth noting that since AF was sampled at Day 68 – well in advance of when spontaneous abortions subsequently occurred – these data suggest that early AF molecular signatures are predictive of later pregnancy loss, rather than reflecting iatrogenic effects of the amniocentesis procedure itself.

Machine-learning based biomarker identification and model validation

Statistically significant values often make poor biomarker predictors, and *vice versa* [141–144]. For example, Zheng *et al*. [143] constructed a model using five genetic variants identified in a genome-wide association study on prostate cancer, concluding that these variants did not enhance predictive power. Similarly, Gränsbo *et al*. demonstrated that while chromosome 9p21 was significantly associated with cardiovascular disease, it did not improve risk prediction [144].

To determine whether this phenomenon also applied to our metabolomic dataset, we used *Random Forest*, an established [145] machine learning algorithm, coupled with model performance evaluation using the ROC curve analysis, to identify potential AF biomarkers of fetal physiology. Indeed, we observed that statistical significance did not align with predictive efficacy. For example, urate levels, which were strongly statistically associated with pregnancy outcome (P=0.0036), produced a relatively poor AUROC value of 0.67. Nonetheless, independent panels consisting of just six AF metabolite relative concentrations confirmed fetal conception method and sex and further predict pregnancy success with relatively high sensitivity and specificity (**Fig. 6**).

537 Recognizing the importance of model validation in independent populations [146], we
538 subsequently analyzed AF samples collected on Day 42 of pregnancy from an independent
539 herd of crossbred beef heifers in Ireland. Initial sPLS-DA revealed clear separation of AF
540 profiles between the two cohorts. This likely reflects a combination of factors, including
541 differences in sampling stage (Day 68 *vs*. Day 42), breed-related genetic variation, maternal
542 characteristics (*e.g*., age and weight; **Supplementary Figures 5A-B**), diet, and environmental
543 and/or management conditions. These influences collectively shape AF composition, producing
544 distinct molecular fingerprints detectable in multivariate analyses.

545 Nonetheless, the same six metabolites used to predict fetal sex in the first cohort (AUROC =
546 0.76), predicted fetal sex in second cohort with a better AUROC of 0.85. These findings confirm
547 that the relative concentrations of just six AF metabolites are efficacious biomarkers of fetal sex
548 across two completely independent groups of cattle. This also engenders confidence in the
549 identified biomarkers for predicting pregnancy success, as the sPLS-DA CV error rate values
550 and AUROC measures for pregnancy outcome prediction are more favorable than for fetal sex
551 prediction.

Study limitations

552 Firstly, although we use the term 'metabolomics', the presented data pertain specifically to the
553 methanol-extractable fraction of the AF metabolome. Moreover, assigning putative compound
554 identifiers (annotation) to peak areas presents significant challenges in semi-quantitative high-
555 throughput untargeted metabolomics analyses [147]. Using standard libraries can lead to
556 misannotations of up to 27.8% [148], primarily due to the highly instrument- and setting-specific
557 nature of metabolite fragmentation spectra [149]. Consequently, conducting targeted
558 quantitative metabolomics against specific standards for select metabolites is important for
559 validating annotations and establishing precise metabolite concentrations.

560 Secondly, while our data hold potential bovine diagnostic relevance, their applicability to human
561 medicine and biological mechanisms requires further investigation. Although there are several
562 similarities between bovine and human pregnancies, notable differences exist. For instance, in
563 cattle, AF begins to accumulate after amnion closure around Day 22 and transitions from
564 predominantly maternal plasma transudate to include increasing fetal urine contribution after
565 renal maturation around Day 40 [150], whereas in humans, amnion closure occurs around Day
566 12 [151]. Also, the bovine placenta is semi-invasive synepitheliochorial, while the human
567 placenta is hemochorial. Such distinctions, among others, highlight that direct extrapolation
568 between bovine and human pregnancy, and *vice versa*, is often inappropriate. Nevertheless,

569 while definitive translational conclusions cannot be drawn, our mention of human pregnancy is
570 intended to place the study in a broader comparative context. Given that many aspects of
571 mammalian development are conserved [152,153], bovine AF analysis may still serve as a
572 valuable resource for nominating candidate biomarkers for subsequent investigation in human
573 studies.

574 Thirdly, the etiologies of pregnancy loss are numerous and complex. Although we identified AF
575 metabolites correlating with spontaneous abortion, our sample size is relatively small. Larger
576 and more sex-balanced cohort studies are essential to enhance the robustness of these data
577 and to assess whether the identified AF metabolites can serve as reliable biomarkers for the
578 early detection of adverse pregnancy outcomes across various pathologies.

579 Fourthly, to increase our sample size for male and female pregnancies, seven pregnancies
580 were generated with sex-sorted semen. Although (a) our conventional and sex-sorted semen
581 originated from the same sire, (b) analysis revealed no correlation between semen source and
582 any measured outcomes, and (c) pregnancy rates following artificial insemination with sex-
583 sorted *vs*. conventional semen are comparable [154,155], altered embryonic gene expression
584 has been reported in bovine embryos produced from sex-sorted *vs*. conventional semen [156].
585 Therefore, this may be considered a confounding factor affecting our results. Lastly, and on a
586 similar note, we compared AI *vs*. IVF derived pregnancies. While effects of AI *vs*. natural service
587 are not apparent [157], this could be considered a potential confounding factor.

Future directions

588 A potentially worthwhile area for further research, not previously discussed, is exploring the
589 influence of the AF microbiome on maternal and fetal physiology. Although this topic is
590 contentious [28,158–160], our data indicate the presence of microbiota-associated compounds
591 in bovine AF. Such inquiries could uncover novel determinants of fetal growth and development.

592 From an applied standpoint, we acknowledge that untargeted MS metabolomics may not be a
593 practical diagnostic modality for routine farm use. A logical next step is to refine these findings
594 into targeted, low-cost, and rapid assays that can be deployed at scale. Such tools would align
595 with emerging precision agriculture approaches, where inexpensive, high-throughput
596 diagnostics have the potential to offset the substantial economic burden of suboptimal
597 reproductive outcomes. Similarly, testing maternal blood, in addition to AF, for biomarkers, as
598 demonstrated in a recent study [72], could provide a less invasive approach to assessing fetal
599 physiology and pregnancy trajectories. Finally, a long-term objective could be to develop AF

transfusions to enhance pregnancy outcomes and improve postnatal well-being in both livestock and humans.

Summary

AF is a valuable yet under-utilized biological resource [161]. This study demonstrates that bovine AF contains reproducible metabolomic biomarkers capable of predicting fetal sex, distinguishing conception method, and forecasting pregnancy outcome with high accuracy. Using machine learning, we identified panels of just six metabolites that achieved strong predictive performance (AUROC values of 0.76-0.91), with the fetal sex model maintaining robust accuracy when validated in an independent cohort of different breed cattle. These findings have immediate practical applications for the livestock industry, where implementation could significantly improve reproductive efficiency and reduce the substantial economic losses associated with early pregnancy failure rates. Beyond agricultural applications, these data could serve as a translational model for human prenatal biomarker development, leveraging the reproductive similarities between species while benefiting from the controlled experimental conditions possible in livestock. The validated framework presented here provides a foundation for developing minimally invasive prenatal diagnostics that could optimize both assisted reproductive technologies and pregnancy management strategies across mammalian species.

DATA AVAILAVBILITY STATEMENT

All data generated or analyzed during this study are included in this published article and corresponding supplementary information files.

CONFLICT OF INTEREST STATEMENT

VAAM, RVS, DCP, VCF, and PJR are employees of ST Genetics Inc. VAAM, PJR, and CAS are co-inventors of a provisional US patent application #63/826,092 based on these data.

AUTHOR CONTRIBUTIONS

Conceptualization: VAAM, PL, PJR, CAS. Collaborative Facilitation: RMR**.** Investigation: VAAM, RVS, DCP, VCF, MM, JMS, MBR. Formal analysis: CAS with support from ID, ZLB, DMF, and AM. Initial manuscript preparation: CAS. Manuscript review and editing: All authors. Funding acquisition and resources: JMS, MBR, AV, XF, PL, PJR, CAS. Supervision: PL, PJR, CAS.

ACKNOWLEDGEMENTS

This work was funded in part by: • Enterprise Ireland Talent Attraction CAM Program (ATI2019-T2-BIO-12966) to JMS (PI) and PL (Co-I); • The European Research Council H2020-MSCA-Individual Fellowship (101021311) to MBR (PI) and PL (Co-I), • The State of Louisiana Board of Regents [LEQSF(2023-26)-RD-A-03] to CAS (PI) and XF (Co-I); • The Audubon Center for Research of Endangered Species (GR-00013800) to CAS (PI) and XF (Co-I); • The Joe W. and Dorothy Dorsett Brown Foundation to CAS (BG008584) (PI); • The United States Department of Agriculture (USDA) Research Capacity (Hatch) funds (LAB-94578) to CAS (PI); and • The Louisiana State University Agricultural Center Collaborative Research Program (PG010315) awarded to CAS (PI), AV (Co-I), and XF (Co-I).

REFERENCES

[1] Bongso TA, Basrur PK. Prenatal diagnosis of sex in cattle by amniocentesis. Vet Rec 1975; 96:124–126.

[2] Vos PL, Pieterse MC, van der Weyden GC, Taverne MA. Bovine fetal fluid collection: Transvaginal, ultrasound-guided puncture technique. Vet Rec 1990; 127:502–504.

[3] Leibo SP, Rall WF. Prenatal diagnosis of sex in bovine fetuses by amniocentesis. Theriogenology 1990; 33:531–552.

[4] Simoni G, Colognato R. The amniotic fluid-derived cells: The biomedical challenge for the third millennium. J Prenat Med 2009; 3:34–36.

[5] Baetz AL, Hubbert WT, Graham CK. Changes of biochemical constituents in bovine fetal fluids with gestational age. Am J Vet Res 1976; 37:1047–1052.

[6] Underwood MA, Gilbert WM, Sherman MP. Amniotic fluid: Not just fetal urine anymore. J Perinatol 2005; 25:341–348.

[7] Michaels J-EA, Dasari S, Pereira L, Reddy AP, Lapidus JA, Lu X, Jacob T, Thomas A, Rodland M, Roberts CT, Gravett MG, Nagalla SR. Comprehensive proteomic analysis of the human amniotic fluid proteome: Gestational age-dependent changes. J Proteome Res 2007; 6:1277–1285.

[8] Bhatti G, Romero R, Gomez-Lopez N, Chaiworapongsa T, Than NG, Theis KR, Galaz J, Gotsch F, Pique-Regi R, Berry SM, Kavdia M, Tarca AL. The amniotic fluid proteome changes with term labor and informs biomarker discovery in maternal plasma. Sci Rep 2023; 13:3136.

[9] Shorey-Kendrick LE, Crosland BA, Spindel ER, McEvoy CT, Wilmarth PA, Reddy AP, Zientek KD, Roberts VHJ, D'Mello RJ, Ryan KS, Olyaei AF, Hagen OL, et al. The amniotic fluid proteome changes across gestation in humans and rhesus macaques. Sci Rep 2023; 13:17039.

[10] Eddleman KA, Malone FD, Sullivan L, Dukes K, Berkowitz RL, Kharbutli Y, Porter TF, Luthy DA, Comstock CH, Saade GR, Klugman S, Dugoff L, et al. Pregnancy loss rates after midtrimester amniocentesis. Obstet Gynecol 2006; 108:1067–1072.

[11] Nizard J. Amniocentesis: Technique and education. Curr Opin Obstet Gynecol 2010; 22:152–154.

[12] Kalogiannidis I, Prapa S, Dagklis T, Karkanaki A, Petousis S, Prapas Y, Prapas N. Amniocentesis-related adverse outcomes according to placental location and risk factors for fetal loss after midtrimester amniocentesis. Clin Exp Obstet Gynecol 2011; 38:239–242.

[13] United States Department of Agriculture (USDA) Economic Research Service (ERS) | The Cattle and beef sector at a glance n.d.

[14] Wiltbank MC, Baez GM, Garcia-Guerra A, Toledo MZ, Monteiro PLJ, Melo LF, Ochoa JC, Santos JEP, Sartori R. Pivotal periods for pregnancy loss during the first trimester of gestation in lactating dairy cows. Theriogenology 2016; 86:239–253.

[15] Kalantari AS, Cabrera VE. Stochastic economic evaluation of dairy farm reproductive performance. Can J Anim Sci 2015; 95:59–70.

[16] Cardoso Consentini CE, Wiltbank MC, Sartori R. Factors that optimize reproductive efficiency in dairy herds with an emphasis on timed artificial insemination programs. Animals 2021; 11:301.

[17] Pohler KG, Reese ST, Franco GA, Oliveira RV, Paiva R, Fernandez L, de Melo G, Vasconcelos JLM, Cooke R, Poole RK. New approaches to diagnose and target reproductive failure in cattle. Anim Reprod 2020; 17:e20200057.

[18] Romero R, Mazaki-Tovi S, Vaisbuch E, Kusanovich JP, Chaiworapongsa T, Gomez R, Nien JK, Yoon BH, Mazor M, Luo J, Banks D, Ryals J, et al. Metabolomics in premature labor: A novel approach to identify patients at risk for preterm delivery. J Matern Fetal Neonatal Med 2010; 23:1344–1359.

[19] Cooper RS. Race in biological and biomedical research. Cold Spring Harb Perspect Med 2013; 3:a008573.

[20] Menon R, Jones J, Gunst PR, Kacerovsky M, Fortunato SJ, Saade GR, Basraon S. Amniotic fluid metabolomic analysis in spontaneous preterm birth. Reprod Sci 2014; 21:791–803.

[21] Leary C, Leese HJ, Sturmey RG. Human embryos from overweight and obese women display phenotypic and metabolic abnormalities. Human Reproduction 2015; 30:122–132.

[22] Baraldi E, Giordano G, Stocchero M, Moschino L, Zaramella P, Tran MR, Carraro S, Romero R, Gervasi MT. Untargeted metabolomic analysis of amniotic fluid in the prediction of preterm delivery and bronchopulmonary dysplasia. PLoS One 2016; 11:e0164211.

[23] Virgiliou C, Gika HG, Witting M, Bletsou AA, Athanasiadis A, Zafrakas M, Thomaidis NS, Raikos N, Makrydimas G, Theodoridis GA. Amniotic fluid and maternal serum metabolic signatures in the second trimester associated with preterm delivery. J Proteome Res 2017; 16:898–910.

[24] Shirasuna K, Iwata H. Effect of aging on the female reproductive function. Contracept Reprod Med 2017; 2:23.

[25] Hallingström M, Barman M, Savolainen O, Viklund F, Kacerovsky M, Brunius C, Jacobsson B. Metabolomic profiles of mid-trimester amniotic fluid are not associated with subsequent spontaneous preterm delivery or gestational duration at delivery. J Matern Fetal Neonatal Med 2022; 35:2054–2062.

[26] Hornaday KK, Wood EM, Slater DM. Is there a maternal blood biomarker that can predict spontaneous preterm birth prior to labour onset? A systematic review. PLoS One 2022; 17:e0265853.

[27] Bourdin-Pintueles A, Galineau L, Nadal-Desbarats L, Dupuy C, Bodard S, Busson J, Lefèvre A, Emond P, Mavel S. Maternal rat metabolomics: Amniotic fluid and placental metabolic profiling workflows. J Proteome Res 2021; 20:3853–3864.

[28] Winters AD, Romero R, Greenberg JM, Galaz J, Shaffer ZD, Garcia-Flores V, Kracht DJ, Gomez-Lopez N, Theis KR. Does the amniotic fluid of mice contain a viable microbiota? Front Immunol 2022; 13:820366.

[29] Carter AM. Animal models of human pregnancy and placentation: alternatives to the mouse. Reproduction 2020; 160:R129–R143.

[30] Ménézo YJR, Hérubel F. Mouse and bovine models for human IVF. Reprod Biomed Online 2002; 4:170–175.

[31] Malhi PS, Adams GP, Singh J. Bovine model for the study of reproductive aging in women: Follicular, luteal, and endocrine characteristics. Biol Reprod 2005; 73:45–53.

[32] Abedal-Majed MA, Cupp AS. Livestock animals to study infertility in women. Anim Front 2019; 9:28–33.

[33] Halstead MM, Ma X, Zhou C, Schultz RM, Ross PJ. Chromatin remodeling in bovine embryos indicates species-specific regulation of genome activation. Nat Commun 2020; 11:4654.

[34] Zhou C, Halstead MM, Bonnet-Garnier A, Schultz RM, Ross PJ. Histone remodeling reflects conserved mechanisms of bovine and human preimplantation development. EMBO Rep 2023; 24:e55726.

[35] Holden SA, Butler ST. Review: Applications and benefits of sexed semen in dairy and beef herds. Animal 2018; 12:s97–s103.

[36] Tiffin GJ, Rieger D, Betteridge KJ, Yadav BR, King WA. Glucose and glutamine metabolism in pre-attachment cattle embryos in relation to sex and stage of development. J Reprod Fertil 1991; 93:125–132.

[37] Ray PF, Conaghan J, Winston RM, Handyside AH. Increased number of cells and metabolic activity in male human preimplantation embryos following in vitro fertilization. J Reprod Fertil 1995; 104:165–171.

[38] Gutiérrez-Adán A, Oter M, Martínez-Madrid B, Pintado B, De La Fuente J. Differential expression of two genes located on the X chromosome between male and female in vitro-produced bovine embryos at the blastocyst stage. Mol Reprod Dev 2000; 55:146–151.

[39] Taylor DM, Handyside AH, Ray PF, Dibb NJ, Winston RM, Ao A. Quantitative measurement of transcript levels throughout human preimplantation development: Analysis of hypoxanthine phosphoribosyl transferase. Mol Hum Reprod 2001; 7:147–154.

[40] Wrenzycki C, Lucas-Hahn A, Herrmann D, Lemme E, Korsawe K, Niemann H. In vitro production and nuclear transfer affect dosage compensation of the X-linked gene transcripts G6PD, PGK, and Xist in preimplantation bovine embryos. Biol Reprod 2002; 66:127–134.

[41] Jiménez A, Madrid-Bury N, Fernández R, Pérez-Garnelo S, Moreira P, Pintado B, de la Fuente J, Gutiérrez-Adán A. Hyperglycemia-induced apoptosis affects sex ratio of bovine and murine preimplantation embryos. Mol Reprod Dev 2003; 65:180–187.

[42] Sturmey RG, Bermejo-Alvarez P, Gutierrez-Adan A, Rizos D, Leese HJ, Lonergan P. Amino acid metabolism of bovine blastocysts: a biomarker of sex and viability. Mol Reprod Dev 2010; 77:285–296.

[43] Kobayashi S, Isotani A, Mise N, Yamamoto M, Fujihara Y, Kaseda K, Nakanishi T, Ikawa M, Hamada H, Abe K, Okabe M. Comparison of gene expression in male and female mouse blastocysts revealed imprinting of the X-linked gene, Rhox5/Pem, at preimplantation stages. Curr Biol 2006; 16:166–172.

[44] Bermejo-Alvarez P, Rizos D, Rath D, Lonergan P, Gutierrez-Adan A. Epigenetic differences between male and female bovine blastocysts produced in vitro. Physiol Genomics 2008; 32:264–272.

[45] Silver MJ, Saffari A, Kessler NJ, Chandak GR, Fall CHD, Issarapu P, Dedaniya A, Betts M, Moore SE, Routledge MN, Herceg Z, Cuenin C, et al. Environmentally sensitive hotspots in the methylome of the early human embryo. Elife 2022; 11:e72031.

[46] Li Y, Sun Y, Zhang X, Wang X, Yang P, Guan X, Wang Y, Zhou X, Hu P, Jiang T, Xu Z. Relationship between amniotic fluid metabolic profile with fetal gender, maternal age, and gestational week. BMC Pregnancy Childbirth 2021; 21:638.

[47] Viana J. Statistics of embryo production and transfer in domestic farm animals: The main trends for the world embryo industry still stand. Embryo Technology Newsletter 2023; 4:20–38.

[48] Fauser BC. Towards the global coverage of a unified registry of IVF outcomes. Reprod Biomed Online 2019; 38:133–137.

[49] Sakka SD, Loutradis D, Kanaka-Gantenbein C, Margeli A, Papastamataki M, Papassotiriou I, Chrousos GP. Absence of insulin resistance and low-grade inflammation despite early metabolic syndrome manifestations in children born after in vitro fertilization. Fertil Steril 2010; 94:1693–1699.

[50] Chen M, Wu L, Zhao J, Wu F, Davies MJ, Wittert GA, Norman RJ, Robker RL, Heilbronn LK. Altered glucose metabolism in mouse and humans conceived by IVF. Diabetes 2014; 63:3189–3198.

[51] Donjacour A, Liu X, Lin W, Simbulan R, Rinaudo PF. In vitro fertilization affects growth and glucose metabolism in a sex-specific manner in an outbred mouse model. Biol Reprod 2014; 90:80.

[52] Niemitz EL, Feinberg AP. Epigenetics and assisted reproductive technology: A call for investigation. Am J Hum Genet 2004; 74:599–609.

[53] Pisarska MD, Chan JL, Lawrenson K, Gonzalez TL, Wang ET. Genetics and epigenetics of infertility and treatments on outcomes. J Clin Endocrinol Metab 2019; 104:1871–1886.

[54] Novakovic B, Lewis S, Halliday J, Kennedy J, Burgner DP, Czajko A, Kim B, Sexton-Oates A, Juonala M, Hammarberg K, Amor DJ, Doyle LW, et al. Assisted reproductive technologies are associated with limited epigenetic variation at birth that largely resolves by adulthood. Nat Commun 2019; 10:3922.

[55] Håberg SE, Page CM, Lee Y, Nustad HE, Magnus MC, Haftorn KL, Carlsen EØ, Denault WRP, Bohlin J, Jugessur A, Magnus P, Gjessing HK, et al. DNA methylation in newborns conceived by assisted reproductive technology. Nat Commun 2022; 13:1896.

[56] Tobi EW, Almqvist C, Hedman A, Andolf E, Holte J, Olofsson JI, Wramsby H, Wramsby M, Pershagen G, Heijmans BT, Iliadou AN. DNA methylation differences at birth after conception through ART. Hum Reprod 2020; 36:248–259.

[57] Mani S, Ghosh J, Coutifaris C, Sapienza C, Mainigi M. Epigenetic changes and assisted reproductive technologies. Epigenetics 2019; 15:12–25.

[58] Mansouri-Attia N, Sandra O, Aubert J, Degrelle S, Everts RE, Giraud-Delville C, Heyman Y, Galio L, Hue I, Yang X, Tian XC, Lewin HA, et al. Endometrium as an early sensor of in vitro embryo manipulation technologies. Proc Natl Acad Sci USA 2009; 106:5687–5692.

[59] Bauersachs S, Ulbrich SE, Zakhartchenko V, Minten M, Reichenbach M, Reichenbach H-D, Blum H, Spencer TE, Wolf E. The endometrium responds differently to cloned versus fertilized embryos. Proc Natl Acad Sci U S A 2009; 106:5681–5686.

[60] Zhou W, Gosch G, Guerra T, Broek D, Wu G, Walker S, Polejaeva I. Amino acid profiles in first trimester amniotic fluids of healthy bovine cloned pregnancies are similar to those of IVF pregnancies, but not nonviable cloned pregnancies. Theriogenology 2014; 81:225–229.

[61] Malin K, Witkowska-Piłaszewicz O, Papis K. The many problems of somatic cell nuclear transfer in reproductive cloning of mammals. Theriogenology 2022; 189:246–254.

[62] Sala RV, Melo LF, Motta JCL, Leffers-Neto L, Carrenho-Sala LC, Fosado M, Moreno JF, Baruselli PS, Wiltbank MC, García-Guerra A. Optimization of a 5-day fixed-time embryo transfer (FTET) protocol in heifers I. Manipulation of circulating progesterone through reutilization of intravaginal progesterone devices during FTET. Theriogenology 2020; 156:171–180.

[63] Absalón-Medina V, Sala R, Bond R. Antimicrobial prophylaxis post-amniocentesis procedures in cattle: A randomized controlled equivalence study. Veterinary and Animal Science 2022; 15:100225.

[64] Pieterse MC, Vos PL, Kruip TA, Wurth YA, van Beneden TH, Willemse AH, Taverne MA. Transvaginal ultrasound guided follicular aspiration of bovine oocytes. Theriogenology 1991; 35:857–862.

[65] Estrada-Cortés E, Jannaman EA, Block J, Amaral TF, Hansen PJ. Programming of postnatal phenotype caused by exposure of cultured embryos from Brahman cattle to colony-stimulating factor 2 and serum. J Anim Sci 2021; 99:skab180.

[66] Tríbulo P, Rivera RM, Ortega Obando MS, Jannaman EA, Hansen PJ. Production and culture of the bovine embryo. Methods Mol Biol 2019; 2006:115–129.

[67] Simintiras CA, Sánchez JM, McDonald M, Martins T, Binelli M, Lonergan P. Biochemical characterization of progesterone-induced alterations in bovine uterine fluid amino acid and carbohydrate composition during the conceptus elongation window. Biology of Reproduction 2018.

[68] Simintiras CA, Sánchez JM, McDonald M, Lonergan P. Progesterone alters the bovine uterine fluid lipidome during the period of elongation. Reproduction 2019; 157:399–411.

[69] Simintiras CA, Sánchez JM, McDonald M, Lonergan P. The influence of progesterone on bovine uterine fluid energy, nucleotide, vitamin, cofactor, peptide, and xenobiotic composition during the conceptus elongation-initiation window. Sci Rep 2019; 9:7716.

[70] Simintiras CA, Sánchez JM, McDonald M, Lonergan P. The biochemistry surrounding bovine conceptus elongation. Biology of Reproduction 2019; 101:328–337.

[71] Rizos D, Ward F, Duffy P, Boland MP, Lonergan P. Consequences of bovine oocyte maturation, fertilization or early embryo development in vitro versus in vivo: Implications for blastocyst yield and blastocyst quality. Mol Reprod Dev 2002; 61:234–248.

[72] Sánchez JM, Gómez-Redondo I, Browne JA, Planells B, Gutiérrez-Adán A, Lonergan P. MicroRNAs in amniotic fluid and maternal blood plasma associated with sex determination and early gonad differentiation in cattle. Biol Reprod 2021; 105:345–358.

[73] Fuhrer T, Heer D, Begemann B, Zamboni N. High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. Anal Chem 2011; 83:7074–7080.

[74] Chen L, Zhernakova DV, Kurilshikov A, Andreu-Sánchez S, Wang D, Augustijn HE, Vich Vila A, Weersma RK, Medema MH, Netea MG, Kuipers F, Wijmenga C, et al. Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. Nat Med 2022; 28:2333–2343.

[75] Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, et al. HMDB 4.0: The human metabolome database for 2018. Nucleic Acids Res 2018; 46:D608–D617.

[76] Pang Z, Zhou G, Ewald J, Chang L, Hacariz O, Basu N, Xia J. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nat Protoc 2022; 17:1735–1761.

[77] Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, Spigelman AF, MacDonald PE, Wishart DS, Li S, Xia J. MetaboAnalyst 6.0: Towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Res 2024; 52:W398–W406.

[78] Hackstadt AJ, Hess AM. Filtering for increased power for microarray data analysis. BMC Bioinformatics 2009; 10:11.

[79] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43:e47.

[80] Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: Testing association with a clinical outcome. Bioinformatics 2004; 20:93–99.

[81] Lu Y, Pang Z, Xia J. Comprehensive investigation of pathway enrichment methods for functional interpretation of LC–MS global metabolomics data. Brief Bioinform 2022; 24:bbac553.

[82] Barker M, Rayens W. Partial least squares for discrimination. Journal of Chemometrics 2003; 17:166–173.

[83] Lê Cao K-A, Boitard S, Besse P. Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. BMC Bioinformatics 2011; 12:253.

[84] Ruiz-Perez D, Guan H, Madhivanan P, Mathee K, Narasimhan G. So you think you can PLS-DA? BMC Bioinformatics 2020; 21:2.

[85] Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: An assessment of 12 outcomes from 8 studies. Trials 2014; 15:139.

[86] Coughlan MP, Rajagopalan KV, Handler P. The role of molybdenum in xanthine oxidase and related enzymes: Reactivity with cyanide, arsenite, and methanol. Journal of Biological Chemistry 1969; 244:2658–2663.

[87] Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. Metabolomics 2013; 9:280–299.

[88] Edmonds CG, Crain PF, Gupta R, Hashizume T, Hocart CH, Kowalak JA, Pomerantz SC, Stetter KO, McCloskey JA. Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). J Bacteriol 1991; 173:3138–3148.

[89] Jühling F, Mörl M, Hartmann RK, Sprinzl M, Stadler PF, Pütz J. tRNAdb 2009: Compilation of tRNA sequences and tRNA genes. Nucleic Acids Res 2009; 37:D159-162.

[90] Alexandrov A, Martzen MR, Phizicky EM. Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. RNA 2002; 8:1253–1266.

[91] Alexandrov A, Chernyakov I, Gu W, Hiley SL, Hughes TR, Grayhack EJ, Phizicky EM. Rapid tRNA decay can result from lack of nonessential modifications. Mol Cell 2006; 21:87–96.

[92] Lin S, Liu Q, Lelyveld VS, Choe J, Szostak JW, Gregory RI. Mettl1/Wdr4-mediated m7G tRNA methylome Is required for normal mRNA translation and embryonic stem cell self-renewal and differentiation. Mol Cell 2018; 71:244-255.e5.

[93] Deng Y, Zhou Z, Ji W, Lin S, Wang M. METTL1-mediated m7G methylation maintains pluripotency in human stem cells and limits mesoderm differentiation and vascular development. Stem Cell Research & Therapy 2020; 11:306.

[94] Ruiz-Arroyo VM, Raj R, Babu K, Onolbaatar O, Roberts PH, Nam Y. Structures and mechanisms of tRNA methylation by METTL1–WDR4. Nature 2023; 613:383–390.

[95] Cheng T, Li X, Chen J, Yang L, Liu J, Song G, Ma H. Investigation of hub genes involved in Turner syndrome using biological informatics methods. Medicine (Baltimore) 2022; 101:e29069.

[96] Han H, Zheng S, Lin S. N7-methylguanosine (m7G) tRNA modification: A novel autophagy modulator in cancer. Autophagy n.d.; 19:360–362.

[97] Muralimanoharan S, Gao X, Weintraub S, Myatt L, Maloyan A. Sexual dimorphism in activation of placental autophagy in obese women with evidence for fetal programming from a placenta-specific mouse model. Autophagy 2016; 12:752–769.

[98] Jeong DS, Lee J-Y, Kim MH, Oh JH. Regulation of sexually dimorphic placental adaptation in LPS exposure-induced intrauterine growth restriction. Molecular Medicine 2023; 29:114.

[99] Ingemarsson I. Gender aspects of preterm birth. BJOG 2003; 110 Suppl 20:34–38.

[100] Di Renzo GC, Rosati A, Sarti RD, Cruciani L, Cutuli AM. Does fetal sex affect pregnancy outcome? Gend Med 2007; 4:19–30.

[101] Orzack SH, Stubblefield JW, Akmaev VR, Colls P, Munné S, Scholl T, Steinsaltz D, Zuckerman JE. The human sex ratio from conception to birth. Proceedings of the National Academy of Sciences 2015; 112:E2102–E2111.

[102] Zhou P, Wang J, Wang J, Liu X. When autophagy meets placenta development and pregnancy complications. Front Cell Dev Biol 2024; 12.

[103] Gonzalez TL, Willson BE, Wang ET, Taylor KD, Novoa A, Swarna A, Ortiz JC, Zeno GJ, Jefferies CA, Lawrenson K, Rotter JI, Chen Y-DI, et al. Sexually dimorphic DNA methylation and gene expression patterns in human first trimester placenta. Biol Sex Differ 2024; 15:63.

[104] Beall MH, van den Wijngaard JPHM, van Gemert MJC, Ross MG. Amniotic fluid water dynamics. Placenta 2007; 28:816–823.

[105] Holtan SG, Creedon DJ, Haluska P, Markovic SN. Cancer and pregnancy: Parallels in growth, invasion, and immune modulation and implications for cancer therapeutic agents. Mayo Clin Proc 2009; 84:985–1000.

[106] Luo Y, Yao Y, Wu P, Zi X, Sun N, He J. The potential role of N7-methylguanosine (m7G) in cancer. J Hematol Oncol 2022; 15:63.

[107] Añazco-Guenkova AM, Miguel-López B, Monteagudo-García Ó, García-Vílchez R, Blanco S. The impact of tRNA modifications on translation in cancer: Identifying novel therapeutic avenues. NAR Cancer 2024; 6:zcae012.

[108] Eriksson JG, Kajantie E, Osmond C, Thornburg K, Barker DJP. Boys live dangerously in the womb. Am J Hum Biol 2010; 22:330–335.

[109] Suzuki T. The expanding world of tRNA modifications and their disease relevance. Nat Rev Mol Cell Biol 2021; 22:375–392.

[110] Katze JR, Farkas WR. A factor in serum and amniotic fluid is a substrate for the tRNA-modifying enzyme tRNA-guanine transferase. Proc Natl Acad Sci U S A 1979; 76:3271–3275.

[111] Cirzi C, Dyckow J, Legrand C, Schott J, Guo W, Perez Hernandez D, Hisaoka M, Parlato R, Pitzer C, van der Hoeven F, Dittmar G, Helm M, et al. Queuosine-tRNA promotes sex-dependent learning and memory formation by maintaining codon-biased translation elongation speed. EMBO J 2023; 42:e112507.

[112] Zhao X, Ma D, Ishiguro K, Saito H, Akichika S, Matsuzawa I, Mito M, Irie T, Ishibashi K, Wakabayashi K, Sakaguchi Y, Yokoyama T, et al. Glycosylated queuosines in tRNAs optimize translational rate and post-embryonic growth. Cell 2023; 186:5517-5535.e24.

[113] Pinborg A, Wennerholm U-B, Bergh C. Long-term outcomes for children conceived by assisted reproductive technology. Fertil Steril 2023; 120:449–456.

[114] Adelman R, Saul RL, Ames BN. Oxidative damage to DNA: Relation to species metabolic rate and life span. Proc Natl Acad Sci U S A 1988; 85:2706–2708.

[115] Cooke MS, Evans MD, Dizdaroglu M, Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. The FASEB Journal 2003; 17:1195–1214.

[116] Cadet J, Davies KJA. Oxidative DNA damage and repair: An introduction. Free Radic Biol Med 2017; 107:2–12.

[117] Kouakou F, Denizot A-L, L'Hostis A, Colet J, Jacques S, Sallem A, Ziyyat A, Vaiman D, Wolf J-P. Plastic used in in vitro fertilization procedures induces massive placental gene expression alterations. EBioMedicine 2023; 91:104572.

[118] Neu J, Afzal A, Pan H, Gallego E, Li N, Li Calzi S, Caballero S, Spoerri PE, Shaw LC, Grant MB. The dipeptide Arg-Gln inhibits retinal neovascularization in the mouse model of oxygen-induced retinopathy. Invest Ophthalmol Vis Sci 2006; 47:3151–3155.

[119] Li N, Ma L, Liu X, Shaw L, Calzi SL, Grant MB, Neu J. Arginyl-glutamine dipeptide or docosahexaenoic acid attenuates hyperoxia-induced small intestinal injury in neonatal mice. J Pediatr Gastroenterol Nutr 2012; 54:499–504.

[120] Fang Y, Yin W, He C, Shen Q, Xu Y, Liu C, Zhou Y, Liu G, Zhao Y, Zhang H, Zhao K. Adverse impact of phthalate and polycyclic aromatic hydrocarbon mixtures on birth outcomes: A metabolome exposome-wide association study. Environmental Pollution 2024; 357:124460.

[121] Taylor A. Aminopeptidases: Structure and function. FASEB J 1993; 7:290–298.

[122] Talabani H, Dreux S, Luton D, Simon-Bouy B, Le Fiblec B, Col J-Y, Guibourdenche J, Oury J-F, Muller F. Fetal anal incontinence evaluated by amniotic fluid digestive enzyme assay in myelomeningocele spina bifida. Pediatr Res 2005; 58:766–770.

[123] Simintiras CA, Drum JN, Liu H, Sofia Ortega M, Spencer TE. Uterine lumen fluid is metabolically semi-autonomous. Commun Biol 2022; 5:191.

[124] Stanton C, Lawn JE, Rahman H, Wilczynska-Ketende K, Hill K. Stillbirth rates: Delivering estimates in 190 countries. The Lancet 2006; 367:1487–1494.

[125] Hille R, Nishino T, Bittner F. Molybdenum enzymes in higher organisms. Coord Chem Rev 2011; 255:1179–1205.

[126] Bonini MG, Miyamoto S, Mascio PD, Augusto O. Production of the carbonate radical anion during xanthine oxidase turnover in the presence of bicarbonate. Journal of Biological Chemistry 2004; 279:51836–51843.

[127] Kayyali US, Donaldson C, Huang H, Abdelnour R, Hassoun PM. Phosphorylation of xanthine dehydrogenase/oxidase in hypoxia. J Biol Chem 2001; 276:14359–14365.

[128] Derrick M, Englof I, Drobyshevsky A, Luo K, Yu L, Tan S. Intrauterine fetal demise can be remote from the inciting insult in an animal model of hypoxia–ischemia. Pediatr Res 2012; 72:154–160.

[129] Kane AD, Camm EJ, Richter HG, Lusby C, Tijsseling D, Kaandorp JJ, Derks JB, Ozanne SE, Giussani DA. Maternal-to-fetal allopurinol transfer and xanthine oxidase suppression in the late gestation pregnant rat. Physiol Rep 2013; 1:e00156.

[130] Boda D, Németh I, Hencz P, Dénes K. Effect of allopurinol treatment in premature infants with idiopathic respiratory distress syndrome. Dev Pharmacol Ther 1984; 7:357–367.

[131] Masaoka N, Nakajima Y, Hayakawa Y, Ohgame S, Hamano S, Nagaishi M, Yamamoto T. Transplacental effects of allopurinol on suppression of oxygen free radical production in chronically instrumented fetal lamb brains during intermittent umbilical cord occlusion. J Matern Fetal Neonatal Med 2005; 18:1–7.

[132] van Dijk AJ, Parvizi N, Taverne M a. M, Fink-Gremmels J. Placental transfer and pharmacokinetics of allopurinol in late pregnant sows and their fetuses. J Vet Pharmacol Ther 2008; 31:489–495.

[133] Torrance HL, Benders MJ, Derks JB, Rademaker CMA, Bos AF, Van Den Berg P, Longini M, Buonocore G, Venegas M, Baquero H, Visser GHA, Van Bel F. Maternal allopurinol during fetal hypoxia lowers cord blood levels of the brain injury marker S-100B. Pediatrics 2009; 124:350–357.

[134] Peeters-Scholte C, Braun K, Koster J, Kops N, Blomgren K, Buonocore G, van Buul-Offers S, Hagberg H, Nicolay K, van Bel F, Groenendaal F. Effects of allopurinol and deferoxamine on reperfusion injury of the brain in newborn piglets after neonatal hypoxia-ischemia. Pediatr Res 2003; 54:516–522.

[135] Benders MJNL, Bos AF, Rademaker CMA, Rijken M, Torrance HL, Groenendaal F, van Bel F. Early postnatal allopurinol does not improve short term outcome after severe birth asphyxia. Arch Dis Child Fetal Neonatal Ed 2006; 91:F163-165.

[136] Chaudhari T, McGuire W. Allopurinol for preventing mortality and morbidity in newborn infants with suspected hypoxic-ischaemic encephalopathy. Cochrane Database Syst Rev 2008:CD006817.

[137] Kaandorp JJ, Benders MJNL, Rademaker CMA, Torrance HL, Oudijk MA, de Haan TR, Bloemenkamp KWM, Rijken M, van Pampus MG, Bos AF, Porath MM, Oetomo SB, et al. Antenatal allopurinol for reduction of birth asphyxia induced brain damage (ALLO-Trial); a randomized double blind placebo controlled multicenter study. BMC Pregnancy Childbirth 2010; 10:8.

[138] Kaandorp JJ, van Bel F, Veen S, Derks JB, Groenendaal F, Rijken M, Roze E, Venema MMAU, Rademaker CMA, Bos AF, Benders MJNL. Long-term neuroprotective effects of allopurinol after moderate perinatal asphyxia: Follow-up of two randomised controlled trials. Arch Dis Child Fetal Neonatal Ed 2012; 97:F162-166.

[139] Ravichandran L, Allen VM, Allen AC, Vincer M, Baskett TF, Woolcott CG. Incidence, intrapartum risk factors, and prognosis of neonatal hypoxic-ischemic encephalopathy among infants born at 35 weeks gestation or more. Journal of Obstetrics and Gynaecology Canada 2020; 42:1489–1497.

[140] Gao T, Zablith NR, Burns DH, Skinner CD, Koski KG. Second trimester amniotic fluid transferrin and uric acid predict infant birth outcomes. Prenat Diagn 2008; 28:810–814.

[141] Lo A, Chernoff H, Zheng T, Lo S-H. Why significant variables aren't automatically good predictors. Proceedings of the National Academy of Sciences 2015; 112:13892–13897.

[142] Jakobsdottir J, Gorin MB, Conley YP, Ferrell RE, Weeks DE. Interpretation of genetic association studies: Markers with replicated highly significant odds ratios may be poor classifiers. PLoS Genet 2009; 5:e1000337.

[143] Zheng SL, Sun J, Wiklund F, Smith S, Stattin P, Li G, Adami H-O, Hsu F-C, Zhu Y, Bälter K, Kader AK, Turner AR, et al. Cumulative association of five genetic variants with prostate cancer. N Engl J Med 2008; 358:910–919.

[144] Gränsbo K, Almgren P, Sjögren M, Smith JG, Engström G, Hedblad B, Melander O. Chromosome 9p21 genetic variation explains 13% of cardiovascular disease incidence but does not improve risk prediction. J Intern Med 2013; 274:233–240.

[145] Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation 2007; 115:654–657.

[146] Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: What, why, how, when and where? Clin Kidney J 2020; 14:49–58.

[147] Kind T, Fiehn O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinformatics 2007; 8:105.

[148] Zhou Z, Luo M, Zhang H, Yin Y, Cai Y, Zhu Z-J. Metabolite annotation from knowns to unknowns through knowledge-guided multi-layer metabolic networking. Nat Commun 2022; 13:6656.

[149] Shen X, Wang R, Xiong X, Yin Y, Cai Y, Ma Z, Liu N, Zhu Z-J. Metabolic reaction network-based recursive metabolite annotation for untargeted metabolomics. Nat Commun 2019; 10:1516.

[150] Li N, Wells DN, Peterson AJ, Lee RSF. Perturbations in the biochemical composition of fetal fluids are apparent in surviving bovine somatic cell nuclear transfer pregnancies in the first half of gestation. Biology of Reproduction 2005; 73:139–148.

[151] Shamsnajafabadi H, Soheili Z-S. Amniotic fluid characteristics and its application in stem cell therapy: A review. Int J Reprod Biomed 2022; 20:627–643.

[152] Mazin PV, Khaitovich P, Cardoso-Moreira M, Kaessmann H. Alternative splicing during mammalian organ development. Nat Genet 2021; 53:925–934.

[153] Gerri C, McCarthy A, Mei Scott G, Regin M, Stamatiadis P, Brumm S, Simon CS, Lee J, Montesinos C, Hassitt C, Hockenhull S, Hampshire D, et al. A conserved role of the Hippo signalling pathway in initiation of the first lineage specification event across mammals. Development 2023; 150:dev201112.

[154] Thomas JM, Locke JWC, Bonacker RC, Knickmeyer ER, Wilson DJ, Vishwanath R, Arnett AM, Smith MF, Patterson DJ. Evaluation of SexedULTRA 4M™ sex-sorted semen in timed artificial insemination programs for mature beef cows. Theriogenology 2019; 123:100–107.

[155] VanWye GM, Andersen CM, Smith EG, Erwin ZL, Spinka C, Poock SE, Thomas JM. Evaluation of later timepoints for fixed-time artificial insemination of beef heifers and cows when using sex-sorted semen. Theriogenology 2024; 214:334–341.

[156] Bermejo-Alvarez P, Rizos D, Rath D, Lonergan P, Gutierrez-Adan A. Sex determines the expression level of one third of the actively expressed genes in bovine blastocysts. Proceedings of the National Academy of Sciences 2010; 107:3394–3399.

[157] Marrella MA, White RR, Dias NW, Timlin C, Pancini S, Currin J, Clark S, Stewart JL, Mercadante VRG, Bradford HL. Comparison of reproductive performance of AI- and natural service-sired beef females under commercial management. Transl Anim Sci 2021; 5:txab114.

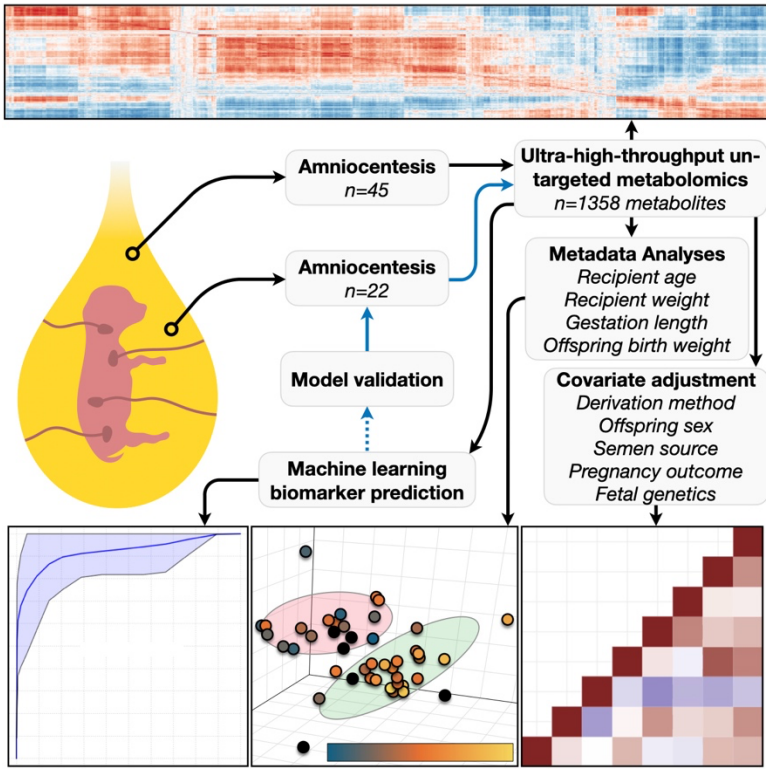
[158] Liu Y, Li X, Zhu B, Zhao H, Ai Q, Tong Y, Qin S, Feng Y, Wang Y, Wang S, Ma J, Yang H. Midtrimester amniotic fluid from healthy pregnancies has no microorganisms using multiple methods of microbiologic inquiry. Am J Obstet Gynecol 2020; 223:248.e1-248.e21.

[159] Stinson LF, Boyce MC, Payne MS, Keelan JA. The not-so-sterile womb: Evidence that the human fetus Is exposed to bacteria prior to birth. Front Microbiol 2019; 10.

[160] Urushiyama D, Suda W, Ohnishi E, Araki R, Kiyoshima C, Kurakazu M, Sanui A, Yotsumoto F, Murata M, Nabeshima K, Yasunaga S, Saito S, et al. Microbiome profile of the amniotic fluid as a predictive biomarker of perinatal outcome. Sci Rep 2017; 7:12171.

[161] Hui L, Bianchi DW. Cell-free fetal nucleic acids in amniotic fluid. Hum Reprod Update 2011; 17:362–371.

FIGURES



Graphical Abstract

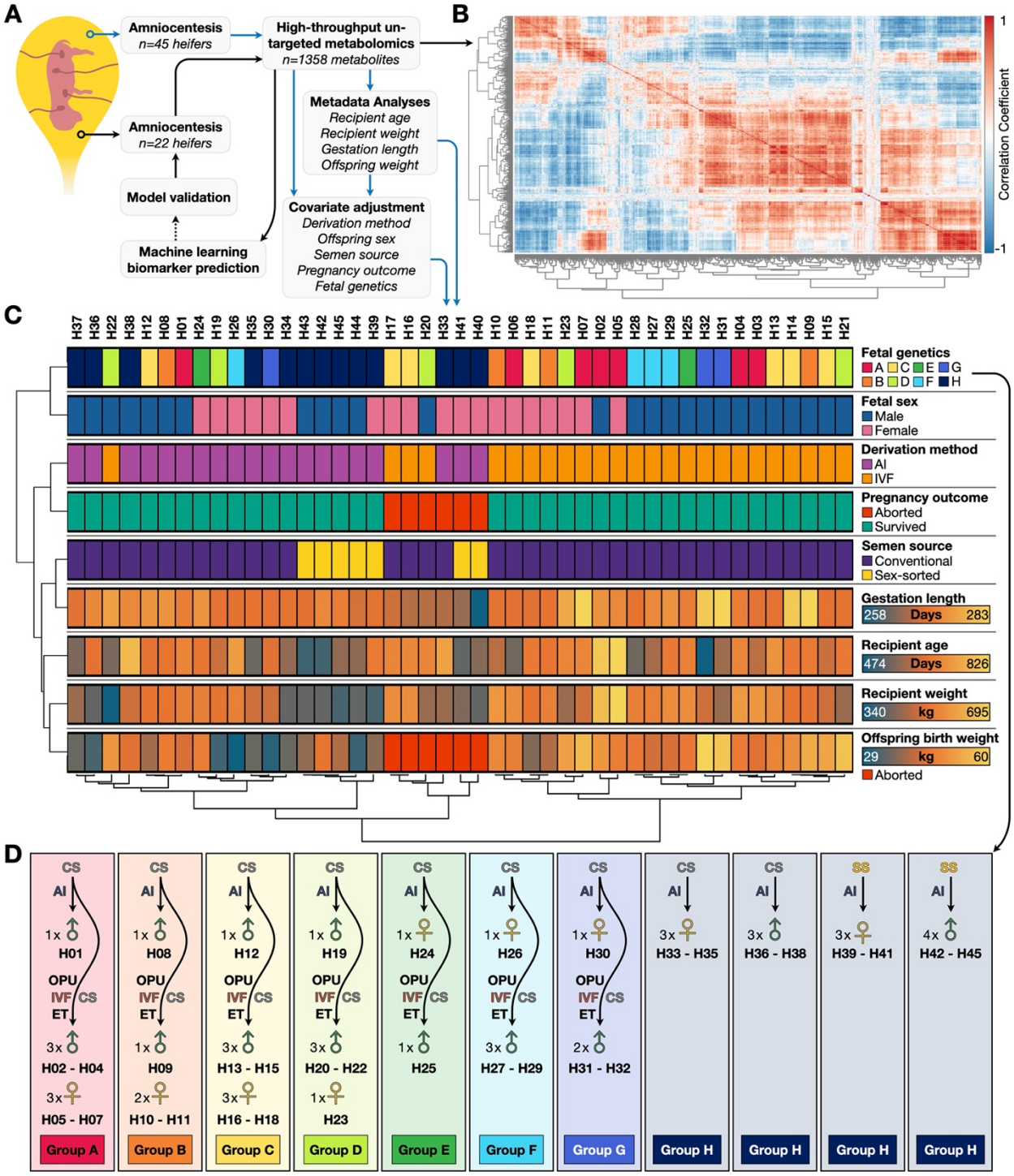

A
Amniocentesis
n=45 heifers
High-throughput un-targeted metabolomics
n=1358 metabolites
Metadata Analyses
Recipient age
Recipient weight
Gestation length
Offspring weight
Amniocentesis
n=22 heifers
Model validation
Machine learning biomarker prediction
Covariate adjustment
Derivation method
Offspring sex
Semen source
Pregnancy outcome
Fetal genetics
B
Correlation Coefficient
1
-1
C
Fetal genetics
A C E G
B D F H
Fetal sex
Male
Female
Derivation method
AI
IVF
Pregnancy outcome
Aborted
Survived
Semen source
Conventional
Sex-sorted
Gestation length
258 Days 283
Recipient age
474 Days 826
Recipient weight
340 kg 695
Offspring birth weight
29 kg 60
Aborted
D
Group A
Group B
Group C
Group D
Group E
Group F
Group G
Group H
Group H
Group H
Group H

Figure 1

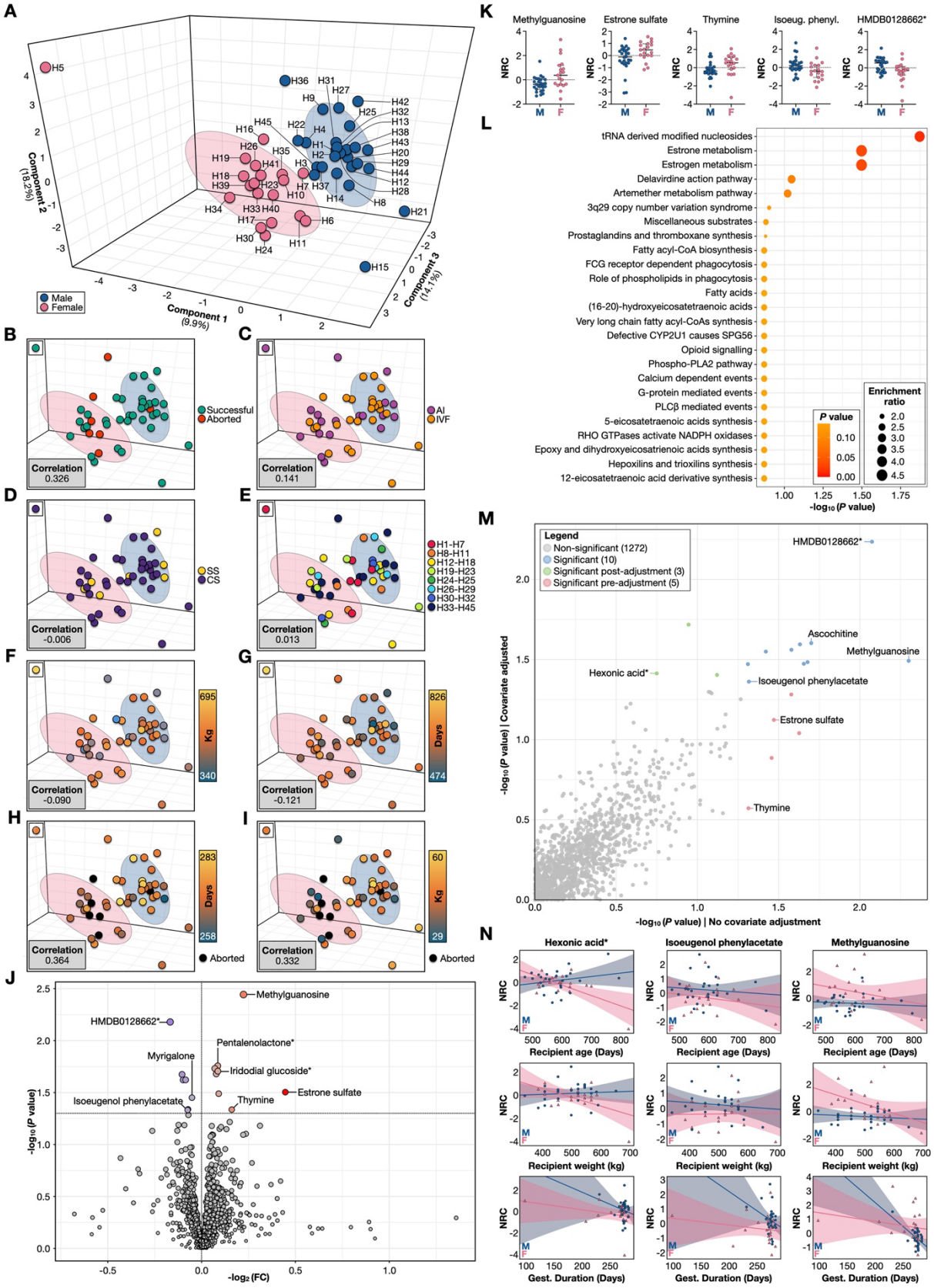

A
Component 1 (9.9%)
Component 2 (18.2%)
Component 3 (14.1%)
Male
Female
B
Successful
Aborted
Correlation 0.326
C
AI
IVF
Correlation 0.141
D
SS
CS
Correlation -0.006
E
H1-H7
H8-H11
H12-H18
H19-H23
H24-H25
H26-H29
H30-H32
H33-H45
Correlation 0.013
F
Kg
Correlation -0.090
G
Days
Correlation -0.121
H
Days
Aborted
Correlation 0.364
I
Kg
Aborted
Correlation 0.332
J
Methylguanosine
HMDB0128662*
Pentalenolactone*
Myrigalone
Iridodial glucoside*
Estrone sulfate
Isoeugenol phenylacetate
Thymine
-log10 (P value)
-log2 (FC)
K
Methylguanosine
Estrone sulfate
Thymine
Isoeug. phenyl.
HMDB0128662*
NRC
M
F
L
tRNA derived modified nucleosides
Estrone metabolism
Estrogen metabolism
Delavirdine action pathway
Artemether metabolism pathway
3q29 copy number variation syndrome
Miscellaneous substrates
Prostaglandins and thromboxane synthesis
Fatty acyl-CoA biosynthesis
FCG receptor dependent phagocytosis
Role of phospholipids in phagocytosis
Fatty acids
(16-20)-hydroxyeicosatetraenoic acids
Very long chain fatty acyl-CoAs synthesis
Defective CYP2U1 causes SPG56
Opioid signalling
Phospho-PLA2 pathway
Calcium dependent events
G-protein mediated events
PLCβ mediated events
5-eicosatetraenoic acids synthesis
RHO GTPases activate NADPH oxidases
Epoxy and dihydroxyeicosatrienoic acids synthesis
Hepoxilins and trioxilins synthesis
12-eicosatetraenoic acid derivative synthesis
P value
Enrichment ratio
-log10 (P value)
M
Legend
Non-significant (1272)
Significant (10)
Significant post-adjustment (3)
Significant pre-adjustment (5)
HMDB0128662*
Ascochitine
Methylguanosine
Hexonic acid*
Isoeugenol phenylacetate
Estrone sulfate
Thymine
-log10 (P value) | Covariate adjusted
-log10 (P value) | No covariate adjustment
N
Hexonic acid*
Isoeugenol phenylacetate
Methylguanosine
Recipient age (Days)
Recipient weight (kg)
Gest. Duration (Days)

Figure 2

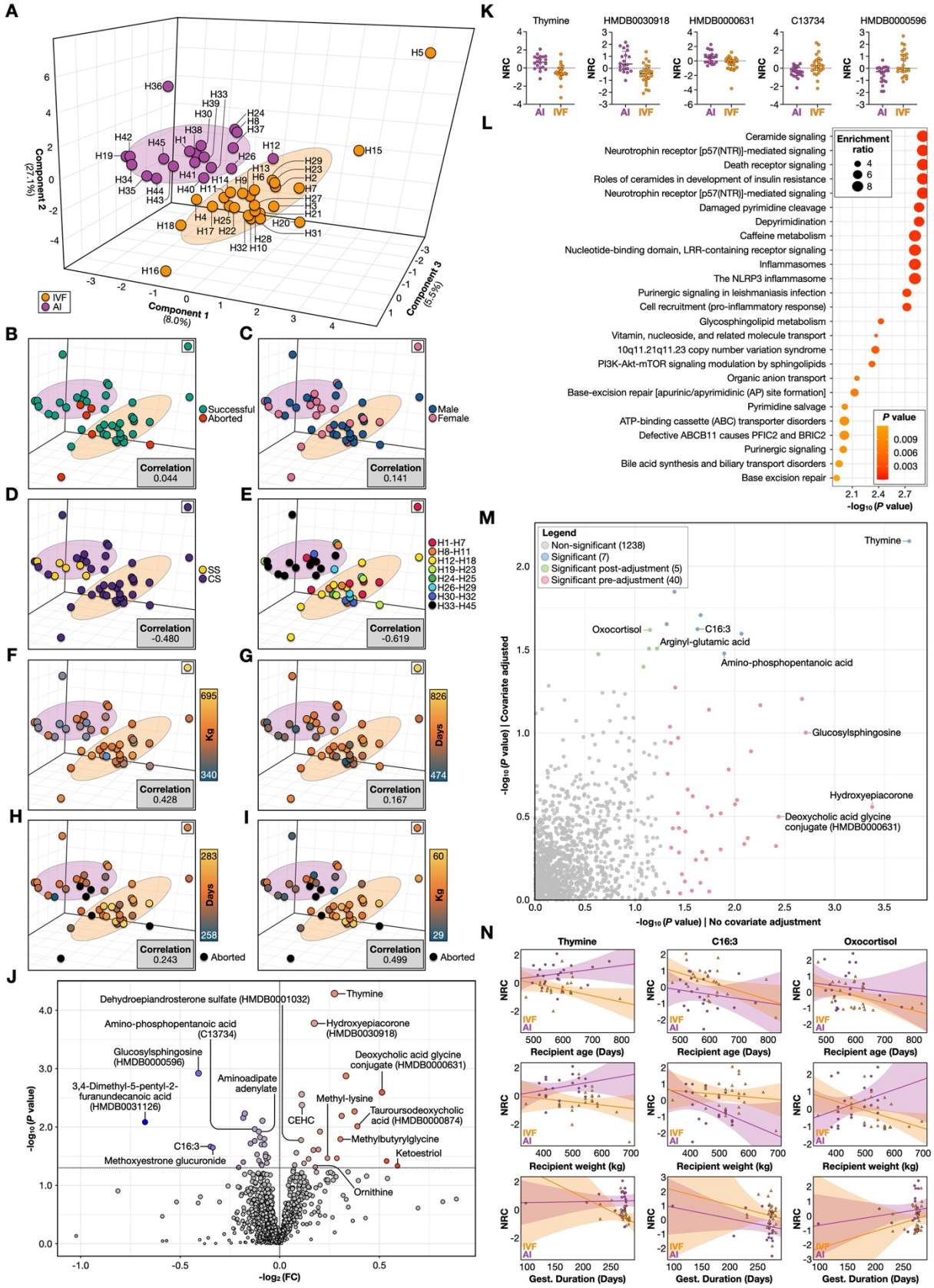

A
Component 1 (8.0%)
Component 2 (27.1%)
Component 3 (5.5%)
IVF
AI
B
Successful
Aborted
Correlation 0.044
C
Male
Female
Correlation 0.141
D
SS
CS
Correlation -0.480
E
H1-H7
H8-H11
H12-H18
H19-H23
H24-H25
H26-H29
H30-H32
H33-H45
Correlation -0.619
F
Kg
Correlation 0.428
G
Days
Correlation 0.167
H
Days
Aborted
Correlation 0.243
I
Kg
Aborted
Correlation 0.499
J
Dehydroepiandrosterone sulfate (HMDB0001032)
Thymine
Amino-phosphopentanoic acid (C13734)
Hydroxyepiacorone (HMDB0030918)
Glucosylsphingosine (HMDB0000596)
Deoxycholic acid glycine conjugate (HMDB0000631)
3,4-Dimethyl-5-pentyl-2-furanundecanoic acid (HMDB0031126)
Aminoadipate adenylate
Methyl-lysine
CEHC
Tauroursodeoxycholic acid (HMDB0000874)
C16:3
Methylbutyrylglycine
Methoxyestrone glucuronide
Ketoestriol
Ornithine
-log10 (P value)
-log2 (FC)
K
Thymine
HMDB0030918
HMDB0000631
C13734
HMDB0000596
NRC
AI
IVF
L
Ceramide signaling
Neurotrophin receptor [p57(NTR)]-mediated signaling
Death receptor signaling
Roles of ceramides in development of insulin resistance
Neurotrophin receptor [p57(NTR)]-mediated signaling
Damaged pyrimidine cleavage
Depyrimidination
Caffeine metabolism
Nucleotide-binding domain, LRR-containing receptor signaling
Inflammasomes
The NLRP3 inflammasome
Purinergic signaling in leishmaniasis infection
Cell recruitment (pro-inflammatory response)
Glycosphingolipid metabolism
Vitamin, nucleoside, and related molecule transport
10q11.21q11.23 copy number variation syndrome
PI3K-Akt-mTOR signaling modulation by sphingolipids
Organic anion transport
Base-excision repair [apurinic/apyrimidinic (AP) site formation]
Pyrimidine salvage
ATP-binding cassette (ABC) transporter disorders
Defective ABCB11 causes PFIC2 and BRIC2
Purinergic signaling
Bile acid synthesis and biliary transport disorders
Base excision repair
Enrichment ratio
P value
-log10 (P value)
M
Legend
Non-significant (1238)
Significant (7)
Significant post-adjustment (5)
Significant pre-adjustment (40)
Thymine
Oxocortisol
C16:3
Arginyl-glutamic acid
Amino-phosphopentanoic acid
Glucosylsphingosine
Hydroxyepiacorone
Deoxycholic acid glycine conjugate (HMDB0000631)
-log10 (P value) | Covariate adjusted
-log10 (P value) | No covariate adjustment
N
Thymine
C16:3
Oxocortisol
NRC
Recipient age (Days)
Recipient weight (kg)
Gest. Duration (Days)

Figure 3

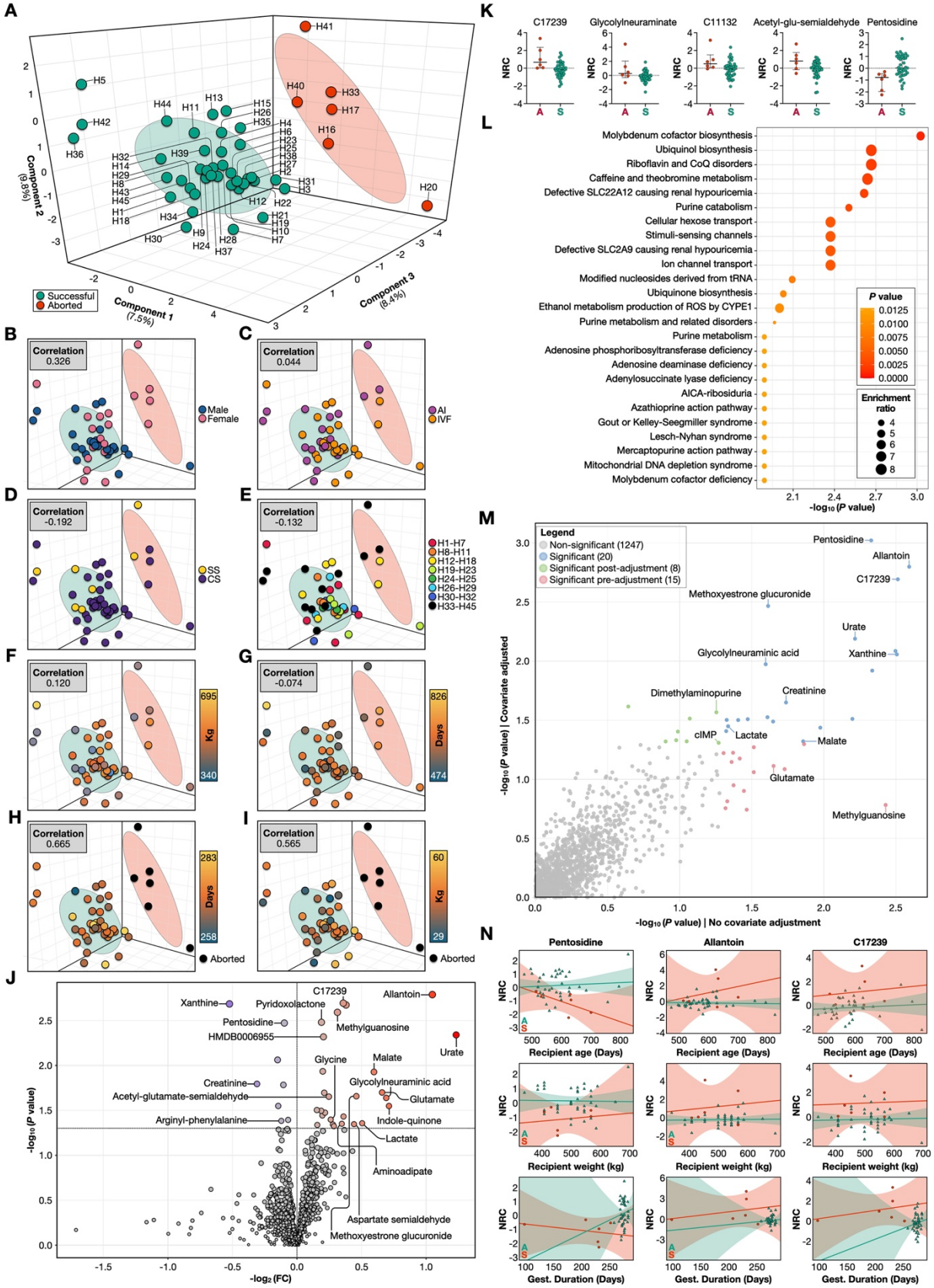

A
Component 1 (7.5%)
Component 2 (9.8%)
Component 3 (8.4%)
Successful
Aborted
B
Correlation 0.326
Male
Female
C
Correlation 0.044
AI
IVF
D
Correlation -0.192
SS
CS
E
Correlation -0.132
H1-H7
H8-H11
H12-H18
H19-H23
H24-H25
H26-H29
H30-H32
H33-H45
F
Correlation 0.120
Kg
G
Correlation -0.074
Days
H
Correlation 0.665
Days
Aborted
I
Correlation 0.565
Kg
Aborted
J
-log2 (FC)
-log10 (P value)
Xanthine
Pyridoxolactone
C17239
Allantoin
Pentosidine
Methylguanosine
HMDB0006955
Urate
Glycine
Malate
Creatinine
Glycolyneuraminic acid
Acetyl-glutamate-semialdehyde
Glutamate
Arginyl-phenylalanine
Indole-quinone
Lactate
Aminoadipate
Aspartate semialdehyde
Methoxyestrone glucuronide
K
C17239
Glycolyneuraminate
C11132
Acetyl-glu-semialdehyde
Pentosidine
NRC
A
S
L
Molybdenum cofactor biosynthesis
Ubiquinol biosynthesis
Riboflavin and CoQ disorders
Caffeine and theobromine metabolism
Defective SLC22A12 causing renal hypouricemia
Purine catabolism
Cellular hexose transport
Stimuli-sensing channels
Defective SLC2A9 causing renal hypouricemia
Ion channel transport
Modified nucleosides derived from tRNA
Ubiquinone biosynthesis
Ethanol metabolism production of ROS by CYPE1
Purine metabolism and related disorders
Purine metabolism
Adenosine phosphoribosyltransferase deficiency
Adenosine deaminase deficiency
Adenylosuccinate lyase deficiency
AICA-ribosiduria
Azathioprine action pathway
Gout or Kelley-Seegmiller syndrome
Lesch-Nyhan syndrome
Mercaptopurine action pathway
Mitochondrial DNA depletion syndrome
Molybdenum cofactor deficiency
P value
Enrichment ratio
-log10 (P value)
M
Legend
Non-significant (1247)
Significant (20)
Significant post-adjustment (8)
Significant pre-adjustment (15)
Pentosidine
Allantoin
C17239
Methoxyestrone glucuronide
Urate
Xanthine
Glycolyneuraminic acid
Dimethylaminopurine
Creatinine
cIMP
Lactate
Malate
Glutamate
Methylguanosine
-log10 (P value) | Covariate adjusted
-log10 (P value) | No covariate adjustment
N
Pentosidine
Allantoin
C17239
Recipient age (Days)
Recipient weight (kg)
Gest. Duration (Days)

Figure 4

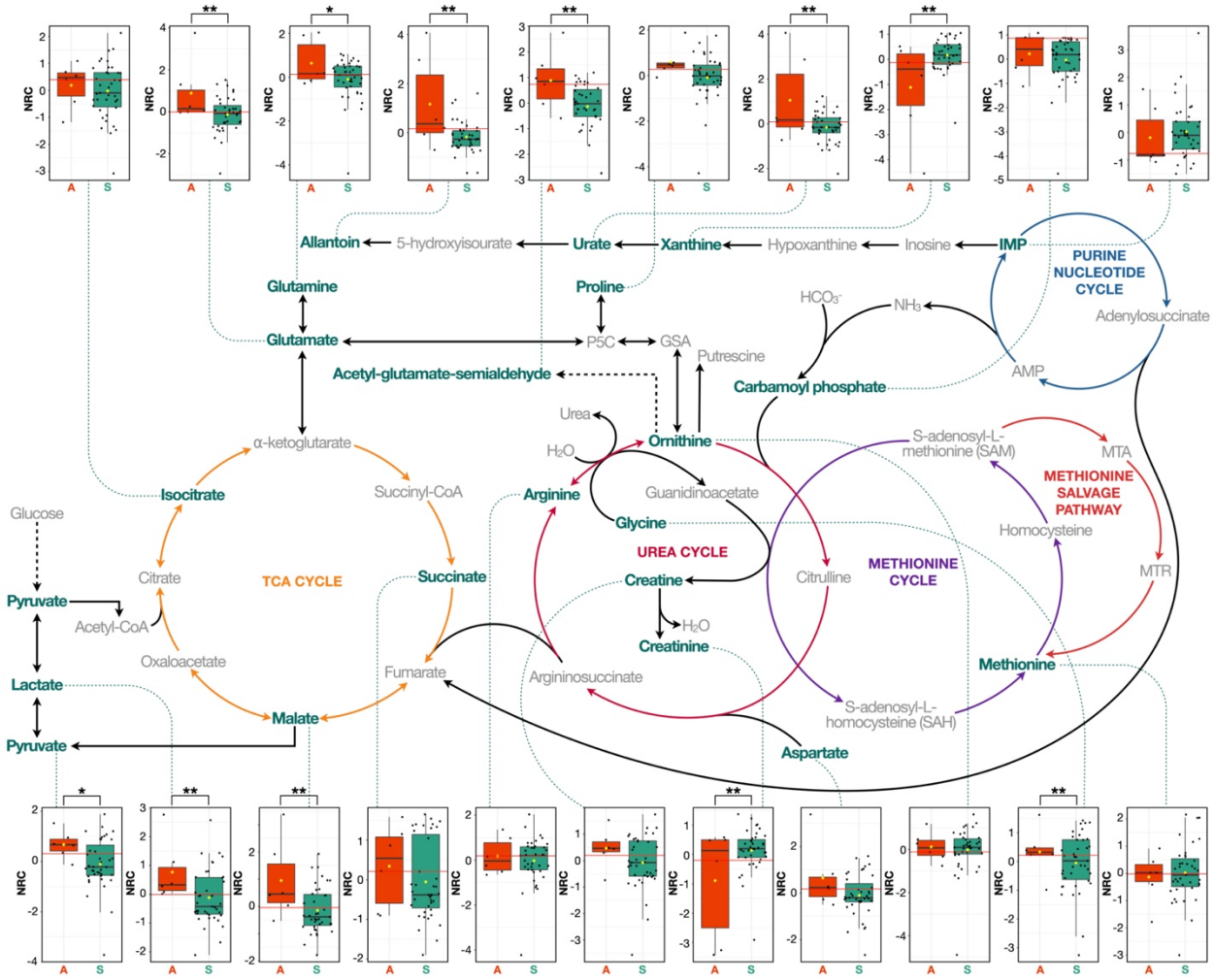

Allantoin
5-hydroxyisourate
Urate
Xanthine
Hypoxanthine
Inosine
IMP
PURINE NUCLEOTIDE CYCLE
Adenylosuccinate
AMP
Glutamine
Proline
Glutamate
P5C
GSA
Putrescine
Acetyl-glutamate-semialdehyde
HCO3-
NH3
Carbamoyl phosphate
Urea
H2O
Ornithine
S-adenosyl-L-methionine (SAM)
MTA
METHIONINE SALVAGE PATHWAY
Homocysteine
MTR
α-ketoglutarate
Isocitrate
Succinyl-CoA
Arginine
Guanidinoacetate
Glycine
UREA CYCLE
Creatine
Citrulline
METHIONINE CYCLE
Glucose
Citrate
TCA CYCLE
Succinate
Pyruvate
Acetyl-CoA
H2O
Creatinine
Oxaloacetate
Fumarate
Argininosuccinate
Methionine
Lactate
Malate
S-adenosyl-L-homocysteine (SAH)
Pyruvate
Aspartate
NRC
A
S

Figure 5

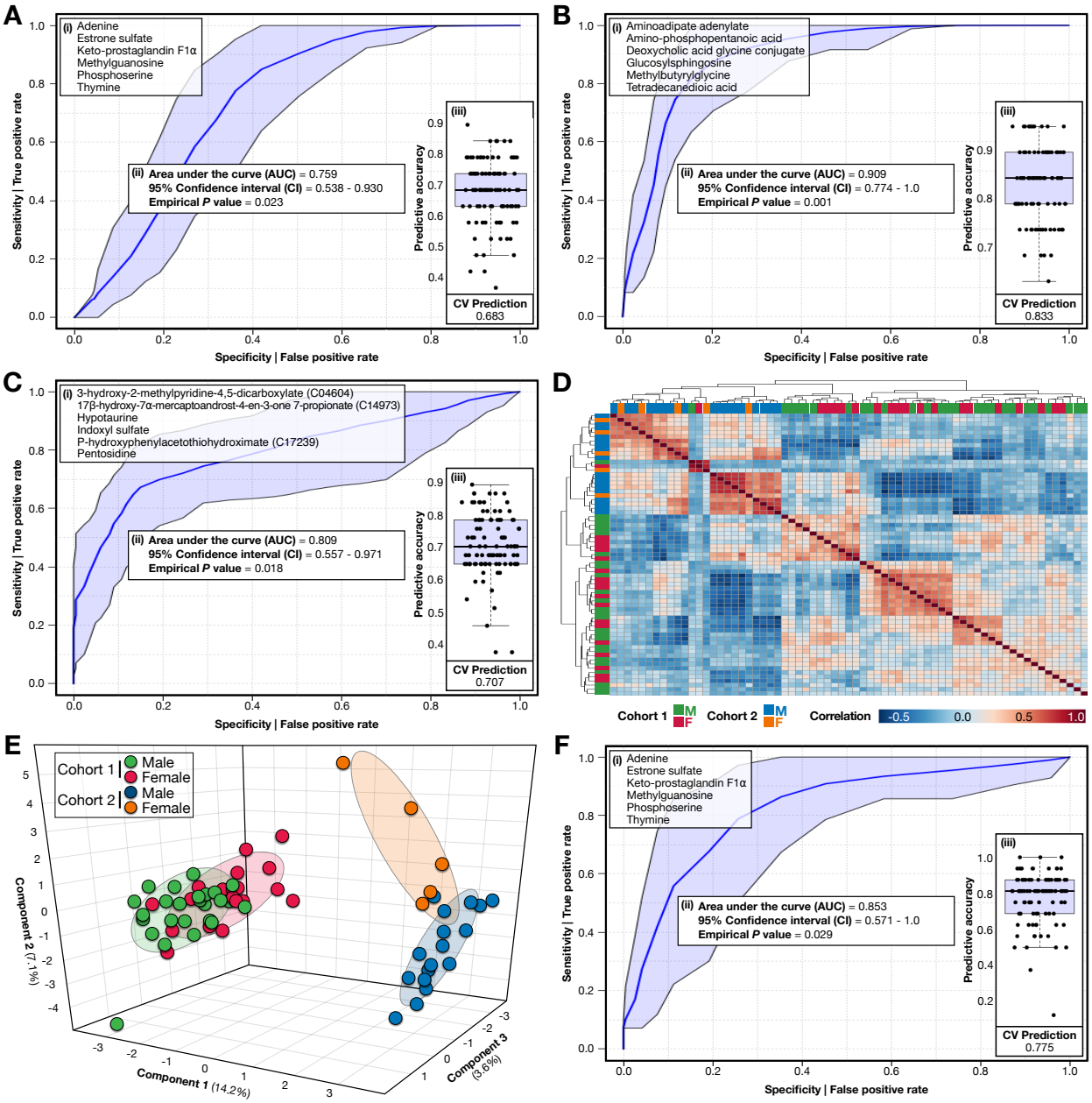

A
(i) Adenine
Estrone sulfate
Keto-prostaglandin F1α
Methylguanosine
Phosphoserine
Thymine
(ii) Area under the curve (AUC) = 0.759
95% Confidence interval (CI) = 0.538 - 0.930
Empirical P value = 0.023
(iii)
CV Prediction
0.683
Sensitivity | True positive rate
Specificity | False positive rate
Predictive accuracy
B
(i) Aminoadipate adenylate
Amino-phosphopentanoic acid
Deoxycholic acid glycine conjugate
Glucosylsphingosine
Methylbutyrylglycine
Tetradecanedioic acid
(ii) Area under the curve (AUC) = 0.909
95% Confidence interval (CI) = 0.774 - 1.0
Empirical P value = 0.001
(iii)
CV Prediction
0.833
C
(i) 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (C04604)
17β-hydroxy-7α-mercaptoandrost-4-en-3-one 7-propionate (C14973)
Hypotaurine
Indoxyl sulfate
P-hydroxyphenylacetothiohydroximate (C17239)
Pentosidine
(ii) Area under the curve (AUC) = 0.809
95% Confidence interval (CI) = 0.557 - 0.971
Empirical P value = 0.018
(iii)
CV Prediction
0.707
D
Cohort 1 M F
Cohort 2 M F
Correlation -0.5 0.0 0.5 1.0
E
Cohort 1 Male Female
Cohort 2 Male Female
Component 1 (14.2%)
Component 2 (7.1%)
Component 3 (3.6%)
F
(i) Adenine
Estrone sulfate
Keto-prostaglandin F1α
Methylguanosine
Phosphoserine
Thymine
(ii) Area under the curve (AUC) = 0.853
95% Confidence interval (CI) = 0.571 - 1.0
Empirical P value = 0.029
(iii)
CV Prediction
0.775

Figure 6

FIGURE LEGENDS

1051 **Figure 1**. *Experimental design summary*. (**A**) Amniotic fluids from Holstein heifers (n=45) were
1052 metabolically profiled. Data were analyzed considering continuous (metadata) and discrete
1053 (covariate) variables. A machine learning approach was then used to predict biomarkers of fetal
1054 sex, conception [*in vitro* fertilization (IVF) *vs*. artificial insemination (AI)], and pregnancy
1055 outcome. Predicted biomarkers of fetal sex were validated using amniotic fluid from an
1056 independent group of heifers (n=22). (**B**) Correlation heatmap of metabolite relative abundance.
1057 (**C**) Metadata and correlation heatmap of amniotic fluid samples from heifers (H) 1 to 45. (**D**)
1058 Schematic depiction of sample generation. Thirteen (13) heifers were artificially inseminated
1059 with conventional semen (CS). Ovum pick-up (OPU) from seven recipients, followed by IVF and
1060 embryo transfer (ET), was performed to generate additional samples (Groups A-G). Finally, 7
1061 heifers were artificially inseminated using sex-sorted (SS) semen. Distinctions between female
1062 (♀) and male (♂) fetuses are also provided.

1063 **Figure 2**. *Amniotic fluid metabolome analysis by fetal sex*. (**A**) Metabolome sparse partial least
1064 squares discriminant analysis (sPLS-DA) of amniotic fluid collected on Day 68 of pregnancy
1065 from heifers gestating a male (blue) *vs*. female (pink) fetus. Individual heifer identifiers (H1-45)
1066 and 95% confidence region ellipses also provided. (**B**-**I**) Metabolome correlation coefficients
1067 between fetal sex and (**B**) pregnancy outcome, (**C**) conception method [*in vitro* fertilization (IVF)
1068 *vs*. artificial insemination (AI)], (**D**) semen source [sexed (SS) *vs*. conventional (CS)], (**E**) fetal
1069 genetics, (**F**) recipient weight, (**G**) recipient age, (**H**) gestation length, and (**I**) offspring birth
1070 weight. Inserts (top left) correspond to sample H5. (**J**) Volcano plot of differentially abundant
1071 ($P \leq 0.05$) metabolites before covariate adjustment with (**K**) boxplots of select corresponding
1072 metabolite normalized relative concentrations (NRC). (**L**) Pathway enrichment analysis of
1073 differentially abundant metabolites. (**M**) Linear model of differentially abundant metabolites
1074 before and after discrete covariate adjustment. (**N**) Scatterplots of select metabolite NRC
1075 against continuous metadata. Asterisks denote predicted metabolites. Additional abbreviation:
1076 3-(6-hydroxy-7-methoxy-2H-1,3-benzodioxol-5-yl)prop-2-enal (HMDB0128662).

1077 **Figure 3**. *Amniotic fluid metabolome analysis by conception method*. (**A**) Metabolome sparse
1078 partial least squares discriminant analysis (sPLS-DA) of amniotic fluid collected on Day 68 of
1079 pregnancy from heifers gestating an *in vitro* fertilization (IVF) *vs*. artificial insemination (AI)
1080 derived fetus. Individual heifer identifiers (H1-45) and 95% confidence region ellipses also
1081 provided. (**B**-**I**) Metabolome correlation coefficients between fetal conception method and (**B**)
1082 pregnancy outcome, (**C**) fetal sex, (**D**) semen source [sexed (SS) *vs*. conventional (CS)], (**E**)

1083 fetal genetics, (**F**) recipient weight, (**G**) recipient age, (**H**) gestation length, and (**I**) offspring birth
1084 weight. Inserts (top right) correspond to sample H5. (**J**) Volcano plot of differentially abundant
1085 ($P \leq 0.05$) metabolites before adjustment with (**K**) boxplots of select corresponding metabolite
1086 normalized relative concentrations (NRC). (**L**) Pathway enrichment analysis of differentially
1087 abundant metabolites. (**M**) Linear model of differentially abundant metabolites before and after
1088 discrete covariate adjustment. (**N**) Scatterplots of select metabolite NRC against continuous
1089 metadata.

1090 **Figure 4**. *Amniotic fluid metabolome analysis by pregnancy outcome*. (**A**) Metabolome sparse
1091 partial least squares discriminant analysis (sPLS-DA) of amniotic fluid collected on Day 68 of
1092 pregnancy from heifers that subsequently delivered successfully (green) or spontaneously
1093 aborted (red). Individual heifer identifiers (H1-45) and 95% confidence region ellipses also
1094 provided. (**B**-**I**) Metabolome correlation coefficients between pregnancy outcome and (**B**) fetal
1095 sex, (**C**) conception method [*in vitro* fertilization (IVF) *vs*. artificial insemination (AI)], (**D**) semen
1096 source [sexed (SS) *vs*. conventional (CS)], (**E**) fetal genetics, (**F**) recipient weight, (**G**) recipient
1097 age, (**H**) gestation length, and (**I**) offspring birth weight. (**J**) Volcano plot of differentially
1098 abundant ($P \leq 0.05$) metabolites before adjustment, with (**K**) boxplots of select corresponding
1099 metabolite normalized relative concentrations (NRC). (**L**) Pathway enrichment analysis of
1100 differentially abundant metabolites. (**M**) Linear model of differentially abundant metabolites
1101 before and after discrete covariate adjustment. (**N**) Scatterplots of select metabolite NRC
1102 against continuous metadata. Asterisks denote predicted metabolites. Additional abbreviations:
1103 3-hydroxy-2-methylpyridine-4,5-dicarboxylate (HMDB0006955),
1104 hydroxyphenylacetothiohydroximate (C17239), and methoxyestrone glucuronide (C11132).

1105 **Figure 5**. *Predicted metabolic pathways underpinning differential fetal metabolism from*
1106 *spontaneously aborted vs. successful pregnancies*. Boxplots of select corresponding metabolite
1107 normalized relative concentrations (NRC) from aborted (A) vs. successful (S) pregnancies also
1108 provided. Asterisks denote significance ($P \leq 0.05$) before (**) and after (*) covariate adjustment.
1109 Additional abbreviations: 5′-methylthioadenosine (MTA), adenosine monophosphate (AMP),
1110 coenzyme A (CoA), glutamate-5-semi-aldehyde (GSA), inosine monophosphate (IMP),
1111 methylthioribose (MTR), pyrroline 5-carboxylate (P5C), and tricarboxylic acid (TCA).

1112 **Figure 6**. *Computational biomarker prediction for fetal physiology and pregnancy success*. (**A**-
1113 **C**) Receiver operating characteristic (ROC) curves generated using the Random Forest
1114 machine learning algorithm by input of 6 metabolites [inserts (i)] to predict (**A**) fetal sex (male vs.
1115 female), (**B**) conception method (*in vitro* fertilization *vs*. artificial insemination), and (**C**)

1116 pregnancy outcome (spontaneously aborted *vs*. successful) for heifer Cohort 1 (n=45).
1117 Corresponding area under the curve (AUC) and empirical *P* values provided [inserts (ii)]. Inserts
1118 (iii) depict corresponding cross-validation (CV) predictive accuracies. (**D**) Hierarchical clustering
1119 heatmap of amniotic fluid profiles from male *vs*. female carrying pregnancies from the initial
1120 (Cohort 1) and model validation (Cohort 2) animals. (**E**) Metabolome sparse partial least
1121 squares discriminant analysis (sPLS-DA) of amniotic fluid from male and female fetuses from
1122 Cohorts 1 and 2 (**F**) ROC curve generated using the same 6 metabolites as in panel (A) to
1123 predict fetal sex in Cohort 2 (n=22).

1124 **Supplementary Figure 1**. *Animal details*. (**A**) Cohort 1A animal synchronization protocol for
1125 artificial insemination (AI). (**B**) Cohort 1B animal synchronization protocol for embryo transfer
1126 (ET). (**C**) Cohort 2 animal synchronization protocol for ET. (**D**) Cohort 1 animal and pregnancy
1127 details, including fetal sex, conception method, semen source, and outcome (discrete covariate
1128 data), in addition to recipient weight (RW), recipient age (RA), gestation length (GL), and
1129 offspring birth weight (BW) (continuous metadata). (**E**) Cohort 2 animal and pregnancy details,
1130 including fetal sex and the day of ET into recipient heifers, in addition to recipient breed, type,
1131 RW, and RA. Additional abbreviations: male (M), female (F), conventional (Conv.), spontaneous
1132 abortion (SA), prostaglandin F2α (PGF2α), gonadotropin releasing hormone (GnRH), and
1133 progesterone-controlled internal drug release device (CIDR), Aberdeen Angus cross (AAX),
1134 Charolais cross (CHX), Limousin cross (LMX), Aberdeen Angus (AA), and Hereford cross
1135 (HEX).

1136 **Supplementary Figure 2**. *Amniotic fluid metabolome by fetal sex*. (**A**) Principal component
1137 analysis (PCA) of amniotic fluid profiles from male (blue) *vs*. female (pink) carrying pregnancies.
1138 (**B**) Hierarchical clustering heatmap of amniotic fluid metabolomic profiles. (**C**) Sparse partial
1139 least squares discriminant analysis (sPLS-DA) cross-validation classification error rates. (**D**)
1140 Targeted pathway impact analysis, integrating pathway enrichment and topology analyses.

1141 **Supplementary Figure 3**. *Amniotic fluid metabolome by conception method*. (**A**) Principal
1142 component analysis (PCA) of amniotic fluid profiles from *in vitro* fertilization (IVF; orange) *vs*.
1143 artificial insemination (AI; purple) derived pregnancies. (**B**) Hierarchical clustering heatmap of
1144 amniotic fluid metabolomic profiles. (**C**) Sparse partial least squares discriminant analysis
1145 (sPLS-DA) cross-validation classification error rates. (**D**) Targeted pathway impact analysis,
1146 integrating pathway enrichment and topology analyses.

1147 **Supplementary Figure 4**. *Amniotic fluid metabolome by pregnancy*. (**A**) Principal component
1148 analysis (PCA) of amniotic fluid profiles from spontaneously aborted (red) *vs*. successful (green)

1149 pregnancies. (**B**) Hierarchical clustering heatmap of amniotic fluid metabolomic profiles. (**C**)
1150 Sparse partial least squares discriminant analysis (sPLS-DA) cross-validation classification error
1151 rates. (**D**) Targeted pathway impact analysis, integrating pathway enrichment and topology
1152 analyses.

1153 **Supplementary Figure 1**. *Animal metadata summary*. (**A**) Cohort 1 animal variables correlation
1154 heatmap.(**B**) Comparison of mean recipient ages between Cohorts 1 and 2 ($P \leq 0.0001$). (**C**)
1155 Comparison of mean recipient weights between Cohorts 1 and 2 ($P \leq 0.0001$).

1156 **Supplementary Table 1**. *Raw metabolomic master dataset*. Internal standard (IS1-IS10),
1157 animal Cohort 1 (H1-45), and Cohort 2 (H46-67) amniotic fluid metabolite peak intensities.