scientific reports

OPEN

Non-invasive giant panda pregnancy and pseudopregnancy biomonitoring by integrated metabolomics and steroidomics

Tom Cools[1,2], Kirsten S. Wilson[3], Desheng Li[4], Catherine Vancsok[5], Baptiste Mulot[6], Antoine Leclerc[6], José Kok[7], Marko Haapakoski[8,9], Mads F. Bertelsen[10], Florian Sicks[11], Franziska Sutter[11], Simon J. Girling[12], Yingmin Zhou[4], Rengui Li[4], Lynn Vanhaecke[1,13]✉ & Jella Wauters[1,2,5]

Understanding the reproductive biology of giant pandas is crucial for their breeding success and conservation. Pregnancy monitoring, however, is challenging due to delayed implantation and obligatory pseudopregnancy, which limits the effectiveness of traditional immunoassays (IA). To remedy this, we combined polar metabolomics and steroidomics to enable a comprehensive view of the urinary molecular composition across six different reproductive phases spanning six pregnant and seven pseudopregnant cycles. Statistical comparisons revealed 696 discriminative features, including 174 features in the early luteal stages, well before the current pregnancy diagnostic window. Pregnant and pseudopregnant cycles showed differences in amino acid, energy, and steroid metabolism before and after CL reactivation, with androgen levels being significantly elevated in pregnant females specifically, suggesting a role in embryo implantation. Interestingly, we detected only one existing IA target metabolite, but identified other discriminative metabolites that may underlie IA signal detection. Finally, we demonstrated that classification models comprising biomarker panels may improve (early) pregnancy diagnosis with accuracies ranging from 0.763 to 1.000 across reproductive phases. These findings offer possibilities for assigning new biomarkers and optimizing IA target selection, thereby enhancing pregnancy monitoring sensitivity and reliability while improving our understanding of giant panda reproductive biology to support conservation efforts.

Keywords Ailuropoda melanoleuca, Metabolomics, Steroidomics, Pregnancy, Endocrine monitoring, UHPLC-HRMS

As one of the most well-known endangered species, the giant panda (*Ailuropoda melanoleuca*) has been a symbol for wildlife conservation for several decades. Their successful conservation program has benefitted greatly from the multitude of studies published on their geographical distribution[1,2] nutrition[3,4] behavior[5,6] genetics[7,8] and reproductive biology[9–11]. Scarce knowledge of the latter was particularly hampering *ex situ* conservation strategies

[1]Laboratory of Integrative Metabolomics, Department of Translational Physiology, Infectiology and Public Health, Faculty of Veterinary Medicine, Ghent University, Salisburylaan 133, Merelbeke 9820, Belgium. [2]Department of Reproduction Biology, Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, 10315 Berlin, Germany. [3]Institute of Regeneration and Repair, Biomolecular and Assay Core, University of Edinburgh, 4-5 Little France Drive, Edinburgh, Scotland, UK. [4]Key Laboratory of SFGA on Conservation Biology of Rare Animals in The Giant Panda National Park, China Conservation and Research Centre for Giant Panda (CCRCGP), Dujiangyan, Sichuan, People's Republic of China. [5]Pairi Daiza Foundation-Pairi Daiza, Domaine de Cambron 1, Brugelette 7940, Belgium. [6]ZooParc de Beauval & Beauval Nature, Avenue du Blanc, Saint-Aignan 41110, France. [7]Ouwehands Dierenpark Rhenen, Grebbeweg 111, Rhenen 3911, The Netherlands. [8]Ähtärin Eläinpuisto OY, Karhunkierros 150, Ähtäri 63700, Finland. [9]Department of Biological and Environmental Science, Konnevesi Research Station, University of Jyväskylä, Sirkkämäentie 220, Konnevesi 44300, Finland. [10]Copenhagen Zoo, Roskildevej 32, Frederiksberg 2000, Denmark. [11]Berlin Zoo, Hardenbergplatz 8, 10787 Berlin, Germany. [12]Royal Zoological Society of Scotland, 134 Corstorphine Road, Edinburgh, Scotland, UK. [13]Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, 19 Chlorine Gardens, Belfast BT9 5DL, Northern Ireland, UK. ✉email: lynn.vanhaecke@ugent.be

early on, as both timing of ovulation and proof of conception were difficult to determine. Although extensive hormonal and behavioral research has greatly assisted in defining the fertile period[6,9,12,13] reliable pregnancy diagnosis is still only possible during the final weeks before birth. This is due to the occurrence of embryonic diapause after conception and because giant pandas experience obligatory pseudopregnancy if no conception occurs after ovulation, even without the act of mating[14,15]. Remarkably, pregnant and pseudopregnant females show highly similar physiological and endocrinological changes during the whole (pseudo)pregnancy, making an early differentiation between the two physiological states challenging[16].

Patterns of urinary progesterone in giant pandas are seemingly identical between pregnancy and pseudopregnancy, with progesterone increasing immediately before ovulation after which it remains slightly elevated for a variable period of 2–3 months, corresponding to the corpus luteum dormancy (CLD) phase[13,17]. Recently, a triphasic pattern was described for this primary progesterone rise, subdivided in CLD1 (variable length), CLD2 (± 80 days prior to birth/end of the cycle), and CLD3 (± 60 days prior to birth/end of the cycle) (Supplementary Fig. S1)[18]. During the CLD phase, the activity of the corpora lutea is still relatively low and the embryo is believed to pause its development. Embryonic development is expected to resume at the start of a secondary, more drastic progesterone rise. Based on this biphasic progesterone profile, giant pandas are believed to experience delayed implantation since the endocrinology corresponds with other species in which embryonic diapause has been established[16,19,20]. Additionally, correlations between the timing of fetal detection with ultrasound and start of the second progesterone peak provided additional evidence that the implantation probably only occurs after the start of the secondary progesterone rise[15]. This secondary rise in progesterone is strongly pronounced and marks the start of the active luteal (AL) phase, lasting for 42 days. In pregnant individuals, this defines the period between blastocyst reactivation and birth of a full-grown, but altricial cub. During the AL phase, estrogens and prostaglandins, more specifically 13,14-dihydro-15-keto-PGF$_{2\alpha}$ (PGFM), have shown diagnostic potential[10,11]. Around the presumed time of implantation, PGFM peaks for the first time, approximately three weeks prior to birth[10]. This peak divides the AL into an early (EAL, pre-PGFM peak) and late (LAL, post-PGFM peak) period (Supplementary Fig. S1)[17]. In pregnant individuals, both estrogens and PGFM will increase from the final two weeks of the LAL onwards, with PGFM climbing to extraordinary high numbers (x20) on average 24 h prior to birth[10,11,17]. Furthermore, glucocorticoids and even progestogens have recently been detected in higher urinary concentrations in pregnant individuals during the (L)AL phase[17]. In contrast, although pseudopregnant females exhibit a similar CL reactivation and AL phase, it is unclear what physiological processes are initiated or maintained during this period, causing these specific hormonal patterns[17].

Despite recent advances in reproductive monitoring of (pseudo)pregnancy in giant pandas, available data on physiological and metabolic alterations during (pseudo)pregnancy remain scarce. Better understanding of these changes would both facilitate the use of artificial reproduction techniques and assist in increasing the success rate of pregnancies[21]. This will be essential to maintain and further improve the genetic diversity within the giant panda population under human care, as certain individual pandas with higher breeding difficulties will be genetically under-represented in the population compared to the pandas that breed more easily. Additionally, understanding hormonal and physiological changes during normal pregnancy can provide the first step in detecting and explaining pregnancy losses, which are believed to occur relatively frequently in this species but remain difficult to detect[22].

This study aimed to obtain a more comprehensive view of the urinary molecular composition of pregnant and pseudopregnant giant pandas over their whole reproductive cycle, starting from the early corpus luteum dormancy phase (CLD1) until the end of the (pseudo)pregnancy (LAL). To this extent, a total of 238 urine samples from six pregnant and seven pseudopregnant cycles were analyzed using ultra-high performance liquid chromatography coupled to high-resolution mass spectrometry (UHPLC-HRMS), combining two complementary approaches: polar metabolomics and steroidomics (Supplementary Fig. S2). By comparing the urinary molecular spectra across different phases of the reproductive cycle, we evaluated the potential of MS-based profiling to discover distinct differences in metabolic changes between pregnancy and pseudopregnancy. The results provide a methodological starting point for future research into physiological differences between the two states and allow us to present novel early pregnancy biomarker (panel) candidates that require validation in larger and independent sample sets.

Methods
Animals
Urine samples of seven female giant pandas were provided by seven zoological institutions. For every sample, 1–15 mL urine was collected and transferred to Eppendorf tubes into 200–1000 μL aliquots during routine hormonal analysis with EIA. In total, urine samples from 13 cycles were included, of which six originated from pregnancies and seven from pseudopregnancies (non-bred cycles) (Table 1). Each cycle was subdivided into six different reproductive phases as described in Wauters et al.[18] i.e., anestrus (start of observations until cross-over between progestogens and estrogens), CLD1 (variable length, starting 7 days after day of peak estrogens), CLD2 (± 80–60 days prior to birth/end of the cycle), CLD3 (± 60 days prior to birth/end of the cycle until early onset of consistent increase in progestogens), EAL (early onset of consistent increase in progestogens until first PGFM peak), and LAL (post first PGFM peak) (Supplementary Fig. S1). Multiple samples were collected longitudinally from the same individuals on different days within the same cycle during every reproductive phase. From these repeated samples, three representative ones were selected per reproductive phase per cycle with sampling times matched, ensuring that each sample corresponded to a similar time point across cycles (Supplementary Table S1). As sample volumes could be insufficient for multiple analyses, selected samples were not always identical for the different performed analyses, i.e., polar metabolomics and steroidomics. The giant pandas were accommodated based on best practice guidelines for animal husbandry as well as recommendations given by the respective supporting Chinese giant panda experts. Because this study only included analysis of urine as a waste product,

Zoological institution	Giant panda stud book (SB) number	Date of birth (dd.mm.yyyy)	Year of sample collection	Cycle type
Pairi Daiza, Belgium	SB741	07.07.2009	2016	Pregnancy
			2019	Pregnancy
Ähtäri Zoo, Finland	SB941	21.09.2014	2019	Pseudopregnancy
			2020	Pseudopregnancy
ZooParc de Beauval, France	SB723	10.08.2008	2017	Pregnancy
			2021	Pregnancy
Edinburgh Zoo, UK	SB569	24.08.2003	2018	Pseudopregnancy
			2020	Pseudopregnancy
Berlin Zoo, Germany	SB868	10.07.2012	2019	Pregnancy
Ouwehands Dierenpark, The Netherlands	SB884	05.08.2013	2018	Pseudopregnancy
			2019	Pseudopregnancy
			2020	Pregnancy
Zoologisk Have København, Denmark	SB919	26.07.2014	2020	Pseudopregnancy

Table 1. Summary of the included animals and cycles.

animal welfare was not affected. Therefore, no specific ethical approval was needed as all animal-related work was conducted according to relevant national and international guidelines. All experimental protocols were permitted by the respective zoos and supporting Chinese giant panda experts.

Sample collection and preparation
All urine samples were collected either via aspiration from the concrete floor with a clean syringe in the absence of the animal or by direct urination in a small container (trained animals). After collection, the samples were transferred to a collection vial and stored at -20 °C until shipment to the Laboratory of Integrative Metabolomics (LIMET) at Ghent University (Belgium). Before analysis, urinary specific gravity (USpG) was measured according to the method described and validated in Wauters et al.[23]. using a handheld digital refractometer designed for cat (PAL-USG (Cat), range 1.000-1.080, Atago™, Fukaya, Japan) or human urine (Pal-10 S (Human), range 1.000-1.060, Atago™, Fukaya, Japan).

Chemicals and reagents
The polar metabolomics and steroidomics methodologies included, respectively, 70 and 56 authentic analytical standards of naturally occurring compounds (Supplementary Table S2). Analytical standards and internal standards (ISTD) were purchased at Sigma Aldrich (St. Louis, Missouri, USA), Steraloids (Newport, Rhode Island, USA), Cayman chemical (Ann Arbor, Michigan, USA), CDN Isotopes (Pointe-Claire, Canada), SERVA Biochemicals (Heidelberg, Germany), VWR Chemicals (Radnor, Pennsylvania, USA), and Santa Cruz Biotechnology (Santa Cruz, California, USA). Solvents were UPLC-MS grade and obtained from Fisher Scientific (Loughborough, UK), Merck (Darmstadt, Germany), Chem-lab (Zedelgem, Belgium), and Biosolve (Valkenswaard, The Netherlands). Ultra-pure water (UPW) was obtained via a purified-water system from Millipore (Brussels, Belgium).

Instrumentation
Metabolomics
For the urinary polar metabolome analysis, the method described by De Paepe et al.[24] was used. Chromatographic separation of the analytes was performed via reversed-phase chromatography with gradient elution using an Acquity HSS T3 C18 column (1.8 mm, 150 mm x 2.1 mm) mounted on a Dionex UltiMate 3000 XRS UHPLC system (Thermo Fisher Scientific, San José, California, USA). Detection of the analytes was performed on a Q-Exactive™ quadrupole-Orbitrap HRMS (Thermo Fisher Scientific, San José, California, USA), which was preceded by heated electrospray ionization (HESI-II source) in polarity switching mode. Prior to analysis, the instrument was calibrated by infusing ready-to-use calibration mixtures (Thermo Fisher Scientific, San José, California, USA). Quality control (QC) samples (constructed by combining 50 μL of the first 20 samples selected at random with a volume higher than 300 μL) were included in the beginning ($n = 10$) and throughout the analysis run ($n = 2$ every 10 samples) to stabilize the system and to correct for potential instrumental drift between batches.

Steroidomics
Steroidomics analysis was executed according to Cools et al.[25]. Analytes were chromatographically separated on a Vanquish Horizon UHPLC system (Thermo Fisher Scientific, San José, California, USA) equipped with a reversed phase Acquity BEH C18 UPLC column (1.7 μm, 2.1×100 mm) (Waters, Millford, Massachusetts, USA). An Orbitrap Exploris 120 HRMS (Thermo Fisher Scientific, San José, California, USA) was used for analyte detection, fitted with an atmospheric-pressure chemical ionization (APCI) source. Initial instrument calibration was first achieved using a Pierce Flexmix calibration solution (Thermo Fisher Scientific, San José, California, USA), which was infused for both ionization modes. QC samples were constructed by combining 80

µL aliquots of a representative subset of 70 samples (assuring that each panda and each reproductive phase was included at least once in both groups) and analyzed identically to the metabolomics analysis described above.

Fragmentation experiments

Fragmentation of significant features was performed for both polar metabolomics and steroidomics on the Q-Exactive™ instrument. For the features selected from the metabolomics method, the same analytical method was largely applied, but this time with parallel reaction monitoring (PRM). Furthermore, the resolution and AGC target were lowered to 17,500 FWHM and 2E + 5, respectively, while the injection time was increased to 100 ms. The isolation window was 1.0 Da and 3 collision energies were applied, namely 15, 30, and 50 eV. For the steroidomics fragmentation, the same strategy was used, but flow rate and gradient were slightly altered (Supplementary Table S3). The initial steroidomics method was also adjusted by adding PRM, and altering the mass resolution, AGC target, and injection time to the same values as the metabolomics fragmentation method. The isolation window was again 1.0 Da and 25, 35, and 50 eV were used as collision energies. Only samples with known high abundances of the selected features were retained. QC samples were constructed by combining 100 µL of each sample and were again analyzed in the beginning ($n = 5$) and throughout ($n = 2$ every 10 or 11 samples) the analysis run. A standard mix containing the authentic analytical standards of the targeted approach was run in the beginning and at the end of the analysis.

Sample extraction

Prior to the analysis of the polar urinary metabolome, 50 µL of urine were diluted with UPW according to the protocol used for human urine by De Paepe et al.[24]. The dilution factor was optimized for giant panda urine to a value of four to obtain the highest number of detected features with good repeatability, i.e., coefficient of variation (CV) of the feature abundance below 30%. In short, samples were first centrifuged for 8 min at 1,000 g and 4 °C followed by the collection of 50 µL of the supernatant in a 1.5-mL Eppendorf tube. Then, 130 µL UPW and 20 µL ISTD solution (10 ng µL^{-1} in UPW) were added. The diluted sample was vortexed for 30 s after which 100 µL was transferred into a glass HPLC vial with insert. Sample extraction prior to the steroidomics analysis was performed with 200 µL of urine according to Cools et al.[25].

Data processing and analysis

Targeted data processing, i.e., only known metabolic features representing the target compounds, was executed using the in-house R package TaPEx or manually in the Quan Browser of XCalibur 4.4 (Thermo Fisher scientific, San José, California, USA) (Supplementary Tables S2)[26]. Untargeted processing, i.e., all detected metabolic features, was carried out with Compound Discoverer (CD) 3.2 (metabolomics) and 3.3 (steroidomics) (Thermo Fisher Scientific, San José, California, USA). Signal intensities (TaPEx) and feature abundances (Quan Browser and CD) were corrected for USpG prior to further data analysis using the adjusted formula of Wauters et al.[23]. Following USpG correction, the untargeted data was explored using principal component analysis (PCA) with iQC (internal quality control samples) normalization to check QC clustering, natural patterning of the samples and identify potential outliers. Potentially interesting features were selected for further data-analysis by building predictive models comparing pregnant and pseudopregnant samples per reproductive phase using orthogonal partial least squares-discriminant analysis (OPLS-DA). These models were validated via the evaluation of specific quality parameters including $Q^2R^2(Y)$, permutation testing, and cross-validated ANOVA (CV-ANOVA) ($p\text{-value} < 0.05$). Univariate statistics comparing pregnant with pseudopregnant sample data were run on the selected untargeted features and the putatively identified targeted compounds using the in-house R pipeline. Depending on their individual normality, which was first tested with a Shapiro-Wilk test ($\alpha = 0.05$), a Welch two-sample t-test (parametric, for normally distributed data) or Wilcoxon rank sum test (non-parametric, for data that was not normally distributed) ($\alpha = 0.05$ for both after adjustment for multiple testing with the Benjamini-Hochberg method[27] was applied. Variables were deemed significantly different between pregnant and pseudopregnant sample data in a certain reproductive phase if their univariate test resulted in an adjusted $p\text{-value} < 0.05$ in that phase while $p\text{-value} > 0.05$ in the anestrus phase (negative control, no correction of the p-value was applied here). MS/MS spectra of significantly different features were further analyzed using isotope and fragmentation pattern analysis with SIRIUS 5.5.7[28], spectral library searching via the Global Natural Product Social Molecular Networking (GNPS)[29] and CD 3.3 for further annotation according to Schymanski et al.[30]. These analyses correspond to Tier 3 confidence, meaning a tentative candidate identification based on computational predictions and spectral similarity was obtained. Wherever possible, the aim was to confirm the identity of discriminative features with the highest level of confidence (Tier 1) by fully matching their accurate mass (*m/z*), retention time, carbon isotope ratio (C12/C13), and fragmentation spectra with those of authentic reference standards analyzed under identical conditions. Finally, in addition to the multi- and univariate data analysis strategies, the predictive potential of biomarker panels was investigated via elastic net regularization ($\alpha = 0.15$) followed by generalized linear modelling (GLM) per reproductive phase[31,32]. All features with CV > 30% in their feature abundance across QC samples were excluded prior to model building. Additionally, all features that were significantly different in the anestrus phase ($p\text{-value} < 0.05$ with univariate statistics) were excluded from all phases to remove inherent variation between groups unrelated to the reproductive state. Optimization of the tuning parameter λ was performed via leave-one-out cross validation (LOOCV), given the small sample set[33].

For correct interpretation of the data analysis, it is important to acknowledge the limited number of individual giant panda bears ($n = 7$) enclosed in this study, even though the majority of mature female giant pandas resident in Europe was included, representing a significant share (~ 1%) of the total global giant panda population under human care. It must also be noted that due to the inclusion of multiple samples from the same individuals, complete independence of all sample groups could not be guaranteed, which may introduce bias and affect variation and model performances. Additionally, confounding factors such as environmental

influences, maturity, diet etc., and their interactions can greatly affect metabolic and hormonal changes across the reproductive cycle. Nevertheless, as the same strategy was applied for all sample groups, the loss of inherent variation within groups is expected to be similar in all sample groups, therefore limiting potential bias. Furthermore, samples from the anestrus phase were included in the study, representing a negative control. Comparing samples from the anestrus phase allowed for correction of potential confounding factors by filtering out variation between sample groups that is likely unrelated to the reproductive state.

Results

Metabolomics analysis

Out of the 70 compounds from the targeted metabolomics list, 43 were detected across the giant panda urine samples (Supplementary Table S4). Univariate statistical tests revealed that USpG corrected peak areas of 15 of those compounds were significantly different between pregnant and pseudopregnant pandas in at least one reproductive phase (Supplementary Table S4). Seven were, however, also significantly different in the anestrus phase. The remaining eight compounds were fragmented leading to three confirmed identities after spectral matching using authentic analytical standards (Tier 1)[30]. Urocanic acid was significantly higher in pseudopregnant urines in CLD3, while two amino acids, L-leucine and L-phenylalanine, were both significantly higher in pregnant samples in the LAL phase (Fig. 1).

The untargeted metabolomics analysis yielded 6,550 and 4,556 features in the positive and negative ionization modes, respectively. Unsupervised data analysis was performed for all samples and per reproductive phase comparing pregnant and pseudopregnant samples (Supplementary Fig. S3A-F). The PCA model exclusively based on samples of the LAL phase showed a better group separation compared to models of the other reproductive phases, with PC1 and PC2 accounting for 22.29% and 12.58% of the variation, respectively (Supplementary Fig. S3F). Some individual sample clustering (pandas SB741 and SB868) was observed in the PCA plot of the CLD3 phase (Supplementary Fig. S3D), which may reflect biological or environmental variation unrelated to pregnancy status (e.g., diet, management, or sample handling conditions across zoos). To further define metabolic differences over the course of the reproductive cycle, supervised OPLS-DA models were constructed for each reproductive phase. When all 11,106 features were included, the OPLS-DA model of the LAL phase showed excellent Q^2 (0.831), $R^2(Y)$ (0.987), permutation testing, and a CV-ANOVA p-value < 0.001. After filtering out all features with a high variance in their feature abundance across QC samples (CV > 30%) or a low
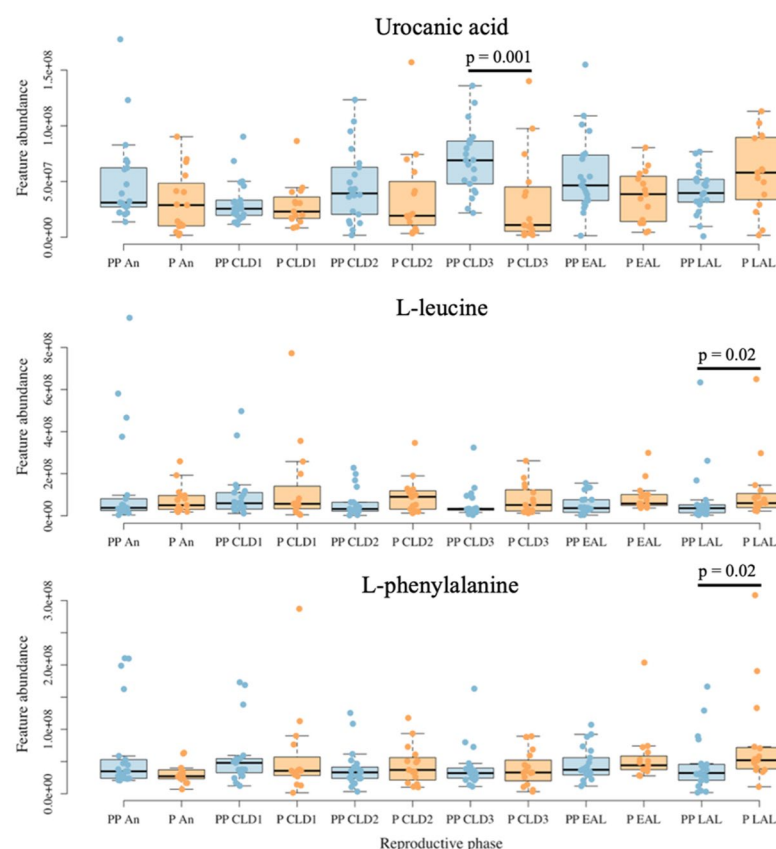

Fig. 1. Targeted polar metabolic feature abundances in pregnant and pseudopregnant giant pandas across the reproductive phases. Feature abundances are based on peak areas of the different samples per reproductive phase in pregnant (P, orange, $n = 15$ per phase) and pseudopregnant (PP, blue, $n = 21$ per phase) cycles. *An* anestrus, *CLD* corpus luteum dormancy phase, *EAL* early active luteal phase, *LAL* late active luteal phase.

VIP score (< 1), new OPLS-DA models were built with more than 2,000 features remaining in all reproductive phases. All newly built models showed excellent $Q^2R^2(Y)$, permutation testing, and CV-ANOVA p-values (Supplementary Table S5). Based on the VIP scores, Jack-knifed confidence interval, and S-plot correlation, 40, 41, 68, 15, and 310 features were retained from the OPLS-DA models constructed with all features for the CLD1, CLD2, CLD3, EAL, and LAL phase, respectively (Supplementary Table S6). After employing univariate statistics, 389 discriminative features were retained across all phases (CLD1: 30, CLD2: 28, CLD3: 40, EAL: 13, LAL: 278) (Supplementary Table S7). Three features could be identified by spectral matching of the fragmentation data with authentic analytical standards (Tier 1)[30]. Kynurenic acid (KYNA) and indole-3-lactic acid (ILA) were both significantly higher in pseudopregnant samples during the CLD2 and CLD3 phase (Fig. 2). In contrast, in the LAL phase these metabolites were significantly higher in the pregnant samples. Tiglylcarnitine decreased in both groups during the early CLD phase compared to anestrus, but returned to higher levels during the late CLD phase (CLD3) in pseudopregnant samples (Fig. 2). The matched spectral data of putatively annotated urinary metabolomics features (UMF) such as UMF1752 and UMF7328 (o-methoxyhippuric acid and an n-aminodecanoic acid, respectively), could not be fully confirmed with analytical standards since they were not commercially available. Preliminary identification was therefore only based on m/z ratio, expected retention time based on known similar compounds, and C12/C13 ratio of the results provided by the fragmentation pattern analysis and spectral library searching (Tier 3)[30]. UMF1752 and UMF7328 were significantly higher in pregnant samples during CLD1 and EAL, respectively (Fig. 2).

Steroidomics analysis

Using steroidomics analysis, 12 out of 56 targeted compounds could be detected in the panda urine (Supplementary Table S8). Seven compounds showed matching fragmentation patterns with the authentic analytical standards, i.e., cortisol, cortisone, tetrahydrocortisone, 1,4-androstadienedione, 11-oxoetiocholanolone, androstenedione, and allopregnanolone sulfate (Tier 1)[30]. The compound putatively identified as 11β-hydroxyetiocholanolone, showed a highly similar fragmentation pattern to the analytical standard of 11β-hydroxyetiocholanolone, but the primary and some minor fragments differed with 2 Da each (Supplementary Fig. S4). This steroid is likely closely related to 11β-hydroxyetiocholanolone with the double bound shifted to another position. It is further referred to as urinary androgen 1 (UA1). UA1 and 11-oxoetiocholanolone were both significantly higher in pregnant urines during the CLD3 and EAL phase (Fig. 3). During the LAL phase, cortisol, 1,4-androstadienedione, and allopregnanolone sulfate were significantly higher in pregnant samples (Fig. 3).

A total of 14,081 features were detected following untargeted steroidomics data processing, 12,474 in the positive and 1,607 in the negative ionization mode. PCA models built per reproductive phase comparing pregnant and pseudopregnant samples revealed no separation between the two groups for any of the reproductive phases (Supplementary Fig. S5). However, some degree of individual sample clustering was observed in the PCA plots potentially reflecting site-specific variation unrelated to reproductive status. Identical to the metabolomics analysis, supervised OPLS-DA models were built per reproductive phase to better investigate the differences between pregnancy and pseudopregnancy. When including all 14,081 features, the OPLS-DA model of the LAL phase again showed excellent Q^2 (0.736), $R^2(Y)$ (0.981), permutation testing, and a CV-ANOVA p-value < 0.001. OPLS-DA models of the other reproductive phases could not be validated. However, when excluding all features with a CV > 30% in their abundance across QC samples and all features with a VIP score < 1, the OPLS-DA model of the CLD3 phase, built with 3,457 features remaining, showed an improved Q^2 (0.562), $R^2(Y)$ (0.995), permutation testing, and a CV-ANOVA p-value of 0.01. Potentially discriminating features were again selected from the OPLS-DA models including all features for CLD1 ($n = 77$), CLD2 ($n = 26$), CLD3 ($n = 30$), EAL ($n = 82$), and LAL ($n = 198$) based on the VIP scores, Jack-knifed confidence interval, and S-plot correlation (Supplementary Table S9). Following univariate statistical tests, the final number of features remaining for CLD1, CLD2, CLD3, EAL, and LAL counted 42, 15, 19, 60, and 171, respectively (Supplementary Table S10). After processing with fragmentation software tools, spectral libraries and CD, the identities of arachidonic acid (AA) and indole could be confirmed based on fragmentation patterns of the corresponding authentic analytical standards (Tier 1)[30]. Both compounds showed significantly higher levels in pregnant compared to pseudopregnant pandas in the LAL phase (Fig. 4). Similarly to the untargeted metabolomics approach, the identity of two other urinary steroidomics features (USF) could not fully be confirmed but they could putatively be annotated as 21-dehydrocortisol (USF2198) and 6,15-diketo-13,14-dihydro-PGF$_{1\alpha}$ (USF5681) based on m/z ratio, expected retention time based on known similar compounds, and C12/C13 ratio of the results provided by the fragmentation pattern analysis and spectral library searching (Tier 3)[30]. Both USF2198 and USF5681 were significantly higher in pregnant samples in the LAL phase (Fig. 4). Moreover, USF5681 already showed significantly higher levels in pregnant samples in both the CLD3 and EAL phase.

Predictive modelling

After QC filtering and removal of the features deemed significantly different in the anestrus phase, 7,486 and 7,397 features remained from the polar metabolomics and steroidomics data, respectively. When combining the metabolomics and steroidomics data, 14,406 features remained for model building. Because all significantly different features were removed, no reliable predictive models could be built with data from the anestrus phase. As expected, the most predictive models were built with data from the LAL phase showing accuracies of 0.972, 1.000, and 1.000 with metabolomics, steroidomics, and the combined data, respectively (Fig. 5). Identical to the PCA plots built with the metabolomics data in this phase, only one sample was wrongly classified as pseudopregnant. All other models, excluding the anestrus phase, however, also showed accuracies higher than 0.760. Furthermore, the models built with only metabolomics data, and with the combination of the metabolomics and steroidomics data for the CLD3 phase offered accuracies higher than 0.940. The number of
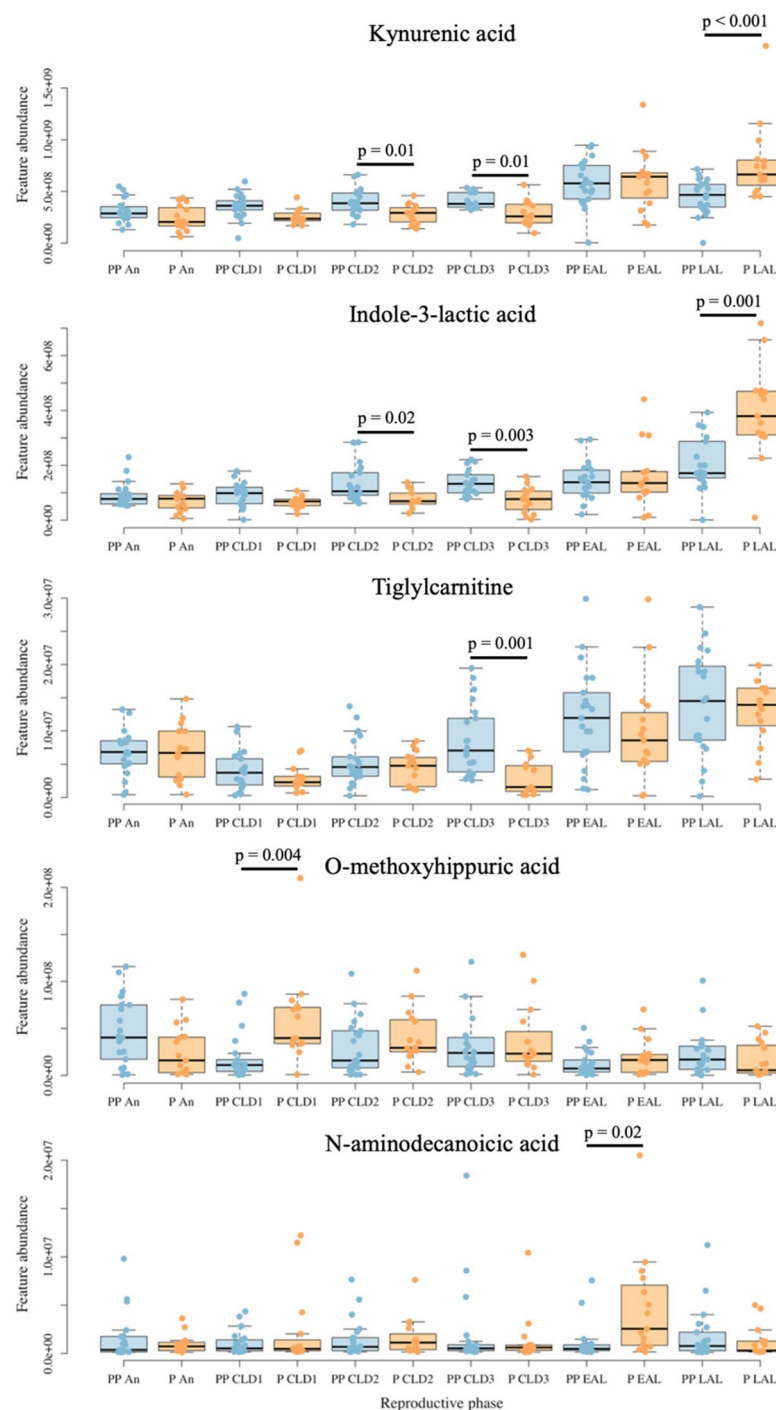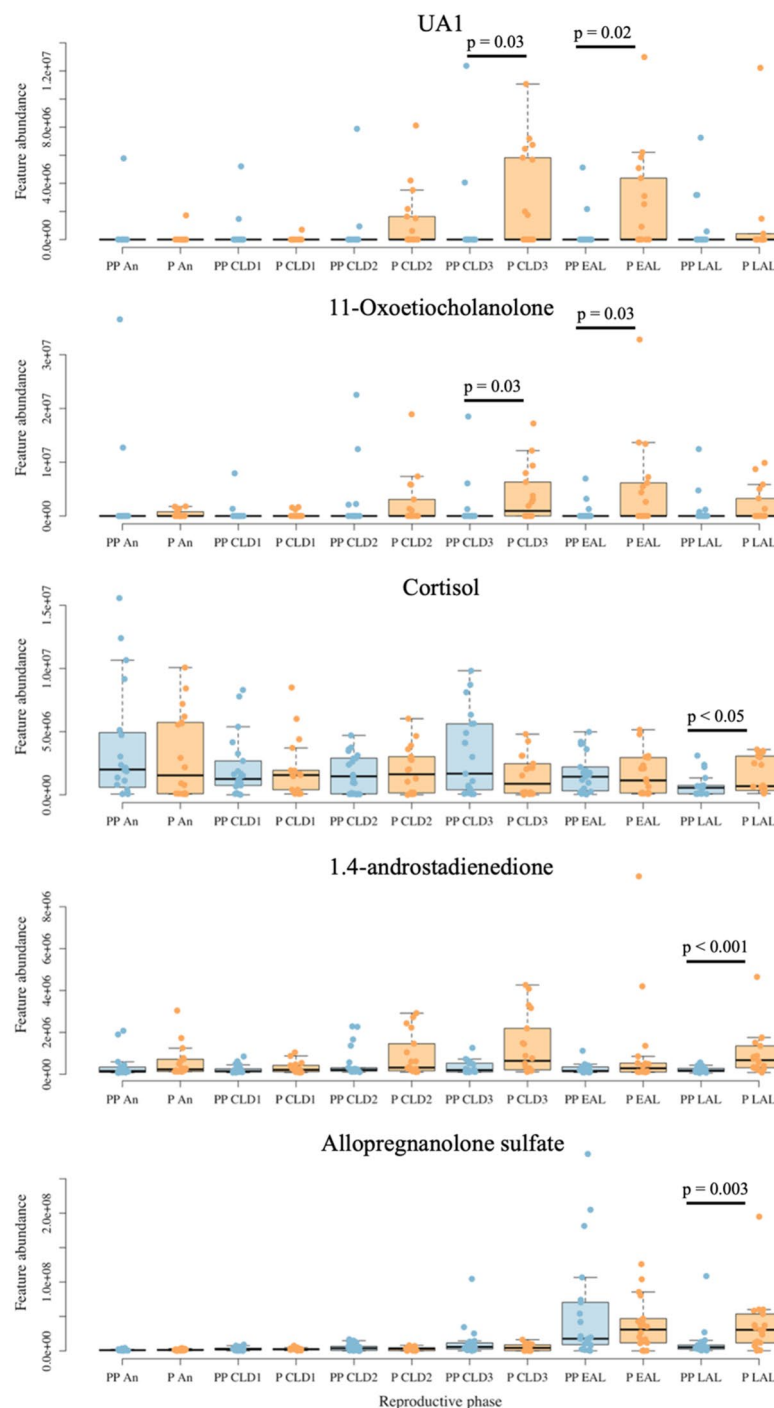
Fig. 2. Untargeted polar metabolic feature abundances in pregnant and pseudopregnant giant pandas across the reproductive phases. Feature abundances are based on peak areas of the different samples per reproductive phase in pregnant (P, orange, $n = 15$ per phase) and pseudopregnant (PP, blue, $n = 21$ per phase) cycles. O-methoxyhippuric acid and n-aminodecanoic acid are putatively identified (Tier 3). *An* anestrus, *CLD* corpus luteum dormancy phase, *EAL* early active luteal phase, *LAL* late active luteal phase.

retained features to build the final models ranged from 64 to 342 (excluding models from the anestrus phase) (Table 2).

Discussion
In this study, we demonstrate distinct urinary molecular profiles between pregnant and pseudopregnant giant pandas by employing an MS-based discovery approach including two different molecular omics methods, i.e., polar metabolomics and steroidomics. Although our findings must be considered within the context of the

Fig. 3. Targeted steroid feature abundances in pregnant and pseudopregnant giant pandas across the reproductive phases. Feature abundances are based on peak areas of the different samples per reproductive phase in pregnant (P, orange, $n = 18$ per phase, except for the CLD1 phase ($n = 17$)) and pseudopregnant (PP, blue, $n = 21$ per phase, except for the anestrus and LAL phase ($n = 20$ and 18, respectively)) cycles. UA1 = unknown androgen 1, an analogue of 11β-hydroxyetiocholanolone. *An* anestrus, *CLD* corpus luteum dormancy phase, *EAL* early active luteal phase, *LAL* late active luteal phase.

relatively small sample size and exploratory nature of this work, the observed differences in metabolic and hormonal features across different reproductive phases provide intriguing insights. These insights can establish a valuable basis for future research aiming at improving pregnancy detection and understanding reproductive physiology in this endangered species. The observed metabolic and hormonal differences were expressed from the early dormancy phase onwards but gained in magnitude towards the late active luteal phase, which is the current window for pregnancy diagnosis through ultrasound and endocrine monitoring. Untargeted polar
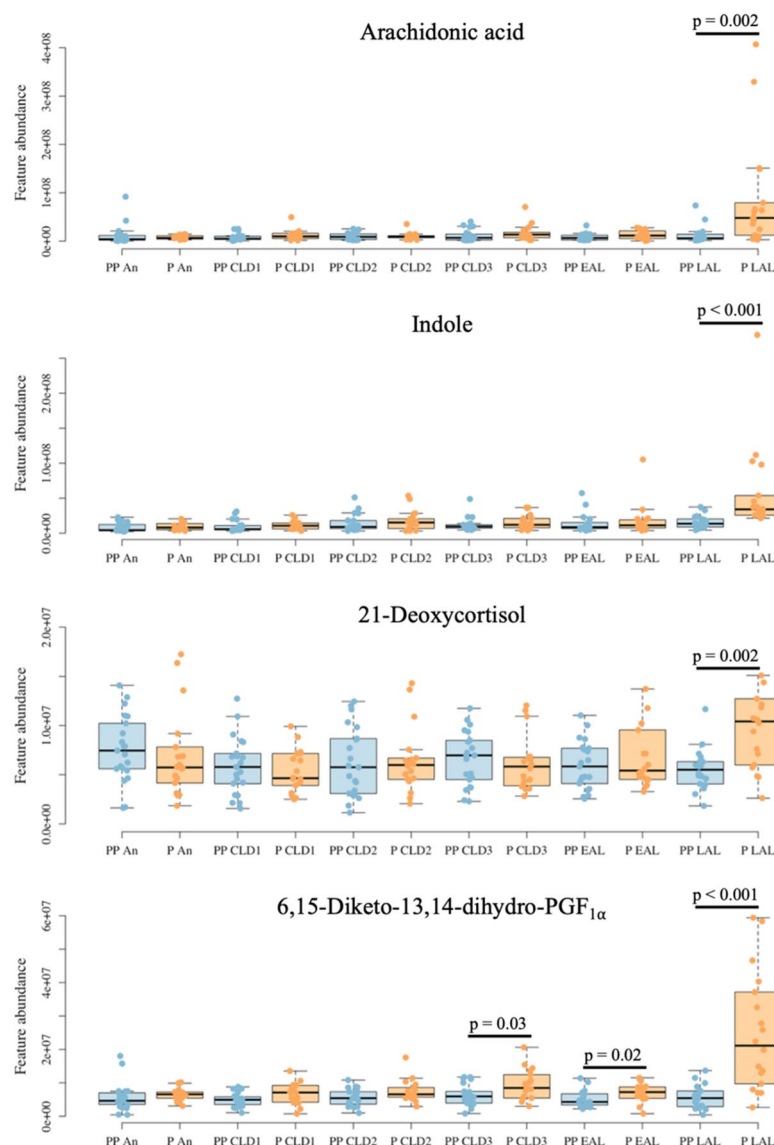
Fig. 4. Untargeted steroid feature abundances in pregnant and pseudopregnant giant pandas across the reproductive phases. Feature abundances are based on peak areas of the different samples per reproductive phase in pregnant (P, orange, $n = 18$ per phase, except for the CLD1 phase ($n = 17$)) and pseudopregnant (PP, blue, $n = 21$ per phase, except for the anestrus and LAL phase ($n = 20$ and 18, respectively)) cycles. 21-deoxycortisol and 6,15-diketo-13,14-dihydro-PGF$_{1\alpha}$ are putatively identified (Tier 3). *An* anestrus, *CLD* corpus luteum dormancy phase, *EAL* early active luteal phase, *LAL* late active luteal phase.

metabolomics and steroidomics analysis revealed a total of, respectively, 389 and 307 significantly different features from which respectively 98 and 76 were retained in the early luteal (CLD) phases. Furthermore, identification of these features revealed changes in various metabolic pathways including amino acid metabolism (urocanic acid, L-leucine, L-phenylalanine, ILA, and KYNA), energy metabolism (tiglylcarnitine and n-aminodecanoic acid), arachidonic acid metabolism (6,15-diketo-13,14-dihydro-PGF$_{1\alpha}$ and AA), and steroid hormone biosynthesis (UA1, 11-oxoetiocholanolone, 1,4-androstadienedione, cortisol, and allopregnanolone sulphate).

Although no features from the CLD1 phase could be fully annotated, UMF1752 was putatively identified as o-methoxyhippuric acid. O-methoxyhippuric acid is not well described in literature, but it could be formed via the conjugation of o-methyl-salicylic acid and glycine[34]. The former is the methylated form of salicylic acid, a compound known for its anti-inflammatory capacities as cyclooxygenase (COX) inhibitor[35,36]. The implantation of an embryo is a pro-inflammatory reaction in which prostaglandins derived via COX play an important role[37]. It is thus probable that a salicylic acid-like compound is increased during the early CLD phase in giant pandas as part of a mechanism to halt implantation of the blastocyst[37].

In the CLD2 and CLD3 phases, two metabolites from the tryptophan pathway, i.e., ILA and KYNA, were identified with levels increased in pseudopregnant urine. In the LAL phase, on the other hand, the levels of both metabolites and a third metabolite from the tryptophan pathway, i.e., indole, were increased in the

Fig. 5. Accuracy dot plots of the GLM models with elastic net regularization per reproductive phase. (**A**) Models built with the metabolomics data. (**B**) Models built with the steroidomics data. (**C**) Models built with the combined metabolomics and steroidomics data. Acc = Accuracy, CI = 95% confidence interval. All models were built after QC filtering (CV > 30%) and exclusion of features significantly different in the anestrus phase ($p < 0.05$). Remaining features used to build the models encompassed 7486 for the metabolomics data, 7,397 for the steroidomics data, and 14,406 for the combined data. *CLD* corpus luteum dormancy phase, *EAL* early active luteal phase, *LAL* late active luteal phase.

pregnant urine instead (Figs. 2 and 4). KYNA has already been described as a potentially important metabolite in giant panda reproduction by Cao et al.[38] who found an increased KYNA excretion in the luteal phase of pregnant pandas compared to their corresponding anestrus phase. The latter finding is also consistent with studies on rats (*Rattus* spp.) and humans[39,40]. KYNA has been hypothesized to be present in higher levels during pregnancy because of its potential neuroprotective effects on the fetus[41]. Alternatively, urinary KYNA levels can be increased because of a higher availability of KYN in the umbilical cord. Indeed, during pregnancy tryptophan is catabolized to protect the fetus against the maternal immune system[42]. As indole and KYN are intermediate products of tryptophan degradation, this process explains the increased indole and KYNA excretion observed during pregnancy[39]. ILA is another metabolite from the tryptophan pathway that has some antioxidant and immunosuppressive effects[43–45]. It can play a role in regulating the microbiome of the newborn cub, which can

Phase	Anestrus	CLD1	CLD2	CLD3	EAL	LAL
Analysis	**Metabolomics (7486 features*)**					
# of samples	36	36	36	36	36	36
(P, PP)	(15, 21)	(15, 21)	(15, 21)	(15, 21)	(15, 21)	(15, 21)
λ	1.000	0.024	0.398	0.030	0.457	1.000
# Of features	19	342	148	316	128	64
Accuracy	0.611	0.889	0.833	0.944	0.861	0.972
(95% CI)	(0.435–0.769)	(0.739–0.969)	(0.672–0.936)	(0.813–0.993)	(0.705–0.953)	(0.855–0.999)
ROC AUC	0.581	0.933	0.902	0.987	0.914	1.000
(95% CI)	(0.383–0.778)	(0.848-1.000)	(0.806–0.997)	(0.960-1.000)	(0.815-1.000)	(1.000–1.000)
Sensitivity	0.200	0.867	0.800	0.867	0.800	1.000
Specificity	0.905	0.905	0.857	1.000	0.905	0.952
Analysis	**Steroidomics (7397 features*)**					
# of samples	38	38	39	39	39	36
(P, PP)	(18, 20)	(17, 21)	(18, 21)	(18, 21)	(18, 21)	(18, 18)
λ	0.060	0.032	0.190	0.263	0.182	0.758
# Of features	399	310	219	170	219	121
Accuracy	0.447	0.763	0.795	0.821	0.821	1.000
(95% CI)	(0.286–0.617)	(0.598–0.886)	(0.635–0.907)	(0.665–0.925)	(0.665–0.925)	(0.903-1.000)
ROC AUC	0.333	0.860	0.823	0.849	0.860	1.000
(95% CI)	(0.157–0.509)	(0.745–0.975)	(0.685–0.961)	(0.713–0.986)	(0.732–0.988)	(1.000–1.000)
Sensitivity	0.611	0.706	0.722	0.778	0.778	1.000
Specificity	0.300	0.810	0.857	0.857	0.857	1.000
Analysis	**Metabolomics and steroidomics (14,406 features*)**					
# of samples	33	35	36	36	34	33
(P, PP)	(13, 20)	(14, 21)	(15, 21)	(15, 21)	(15, 19)	(15, 18)
λ	1.000	0.302	0.251	0.120	0.079	1.000
# Of features	49	209	235	261	317	102
Accuracy	0.545	0.857	0.861	0.972	0.853	1.000
(95% CI)	(0.364–0.719)	(0.697–0.952)	(0.705–0.953)	(0.855–0.999)	(0.689–0.950)	(0.894-1.000)
ROC AUC	0.331	0.929	0.902	0.952	0.947	1.000
(95% CI)	(0.135–0.526)	(0.847-1.000)	(0.801-1.000)	(0.858-1.000)	(0.882-1.000)	(1.000–1.000)
Sensitivity	0.000	0.714	0.800	0.933	0.800	1.000
Specificity	0.900	0.952	0.905	1.000	0.895	1.000

Table 2. Predictive performance parameters of the GLM models with elastic net regularization per reproductive phase. P = pregnant, PP = pseudopregnant, ROC = receiving operator curve, AUC = area under the curve, CI = confidence interval. λ was optimized per reproductive phase via LOOCV. The # of features corresponds to the total number of variables retained in the final model. *Number of features remaining after QC filtering (CV > 30%) and exclusion of features significantly different in the anestrus phase ($p < 0.05$).

explain the higher levels in pregnant pandas during the LAL phase[43–45]. The reasons why KYNA and ILA are both higher in pseudopregnant pandas during the CLD2 and CLD3 phases remain unclear.

As a short-chain acylcarnitine[46] tiglylcarnitine functions as a fatty acid transporter into the mitochondria for energy production via fatty acid oxidation[47]. During specific conditions with increased energy demand (e.g., exercise or pregnancy) or lower energy intake (e.g., fasting), fatty acid oxidation will increase and thus also the production of acylcarnitines[48,49]. In pregnant pandas, tiglylcarnitine levels indeed start to increase during the EAL phase. This is consistent with results regarding the fecal output of giant pandas published by Wauters et al.[17] where a steep decline was noted midway through the EAL phase, approximately 30 days before parturition and thus around the same time window as the EAL samples of this study. However, further investigations are warranted to reveal why tiglylcarnitine levels appear to increase earlier (during CLD3) in pseudopregnant pandas (Fig. 2). UMF7328 in its turn, was increased in pregnant urine samples from the EAL phase and although its identity could not be fully confirmed, it was putatively classified as an n-aminodecanoic acid, which is a medium-chain fatty acid (MCFA). MCFAs are not only readily available energy sources, they also play an important part in fetal development and specifically reduce the chances on early pregnancy loss by improving ovarian steroidogenesis and endometrial receptivity[50–52]. The EAL phase marks the implantation/attachment period of the giant panda embryo and is thus a crucial period that could benefit from these protective effects.

In the final phase of the reproductive cycle, i.e., the LAL phase, several of the observed metabolic and hormonal differences were related to AA metabolism and steroid hormone biosynthesis. Both have been extensively investigated in giant pandas via the use of enzyme immunoassays (EIA)[10,11,16,53,54]. Nevertheless, confirmation of

the identity of the detected hormones is absent when merely using EIA for non-invasive endocrine monitoring as the hormone metabolization process in giant pandas is yet to be unraveled. For example, prostaglandin levels diverge between pregnant and pseudopregnant pandas during the final two weeks to prepare the pregnant female for parturition and are therefore monitored via an EIA that targets PGFM, a metabolite of PGF$_{2\alpha}$[10,17,55]. According to our results, however, giant pandas do not excrete PGFM. Instead, another metabolite is excreted during these stages, namely 6,15-diketo-13,14-dihydro-PGF$_{1\alpha}$, which is probably picked up via EIA cross-reactivity. Remarkably, 6,15-diketo-13,14-dihydro-PGF$_{1\alpha}$ already shows significant differences in earlier stages than can be seen with PGFM EIA monitoring (Fig. 4). More specific EIAs should therefore be developed as they could possibly also pick up this trend and thus provide earlier signs for pregnancy in giant pandas than the current PGFM EIA. Similar to PGFM, we were not able to detect progesterone in any of the urine samples, despite this usually being the main target of the progesterone EIA used for giant panda reproductive monitoring[16,17]. The detection of allopregnanolone sulfate in giant panda urine on the other hand, has not been described in literature yet, while our results suggest it is the main urinary progestogen in this species. In contrast, the presence of cortisol, which has been targeted with EIA in giant pandas, was confirmed with our analysis. It also demonstrated the expected increase during the LAL phase, priming both the fetus and the mother for the nearing parturition[17,54,56,57]. However, a second potential glucocorticoid, putatively identified as 21-dehydrocortisol, showed a more distinct increase than cortisol towards the end of the cycle in pregnant females which was absent in their pseudopregnant counterparts (Fig. 4). As both steroids show highly similar structures, it is possible that 21-dehydrocortisol is detected with the same EIA as cortisol. Although glucocorticoids stimulate the synthesis of estrogens during these final weeks, resulting in improved myometrial activity, this could not be observed with our method as we were unable to identify any estrogens yet. Nevertheless, estrogens are synthetized via the conversion of androgens by aromatase enzymes[58]. As the steroid biosynthesis pathway gradually alters to favor estrogen over progestogen production, androgen levels can also increase being an intermediate between progestogens and estrogens. This can explain the higher levels of 1,4-androstadienedione observed solely in the pregnant giant pandas in the LAL phase in this study. Remarkably, increased levels of two other androgens, UA1 and 11-oxoetiocholanolone, were also noted in earlier stages of pregnant pandas in this study, i.e., the CLD3 and EAL phase (Fig. 3). The EAL corresponds with the CL reactivation and the window for embryo implantation. Androgens play a role in the normal endometrial physiology for embryo implantation both via the conversion to estrogens as via direct binding with androgen receptors[59,60]. It could be hypothesized that in giant pandas, these androgens are involved in preparing the endometrium for implantation of the embryo around reactivation of the CL. As the levels are higher in pregnant females, embryonic signals probably assist in their upregulation.

Despite the successful annotation of multiple features in this study, identification of biomarkers remains a bottleneck in untargeted metabolomics analyses. Therefore, it was explored whether panels of metabolic (including steroids) features could be used to classify samples using binomial logistic regression. Given the small set of samples ($n = 33–39$ per phase), the high number of variables ($n = 7,397 – 14,406$), and the absence of a set of independent samples for model validation, a regularization penalty was applied to prevent overfitting of the models on the training data. Elastic net regularization was selected because it allows feature selection by removing unimportant variables from the model, while taking into account potential correlations between variables[31,32]. The latter is expected to be present as most features will be part of biological systems, which often show complex interactions. The excellent parameters of the logistic regression models, particularly those built from data of the early stages, i.e., CLD (Table 2), suggest that panels of biomarkers measured with UHPLC-HRMS could provide an alternative strategy to the use of EIA for early pregnancy diagnosis. Nevertheless, it must be noted that these models were only validated on the training data. They will need to be validated on newly collected data to fully confirm their potential as diagnostic tools.

Conclusions

To the best of our knowledge, this is the first study to employ a non-invasive metabolomics approach to investigate metabolic and hormonal differences between pregnant and non-pregnant (e.g., proven pseudopregnancies) giant pandas across different phases of their reproductive cycles. Remarkably, our results demonstrate that discriminative changes can be observed as early as the start of the embryonic diapause, i.e., CLD1, long before the current window of pregnancy diagnosis. Our findings did not only confirm differences between pregnant and pseudopregnant pandas in prostaglandin and steroid hormone biosynthesis, but also provided new insights in potential alterations of amino acid and energy metabolism during the reproductive cycle. After further elucidation and validation of the underlying mechanisms of these changes in an independent sample set, these discriminating metabolites could be explored as possible biomarkers for pregnancy monitoring in giant pandas. Alternatively, the potential of biomarker panels, without the need for individual feature annotation, was also demonstrated. Similar to the individually identified metabolites, these classification models will require additional validation in independent sample sets before implementation can be considered. Furthermore, for the first time to our knowledge, we were able to fully annotate discriminatory urinary steroids, some of which are already included in non-invasive pregnancy monitoring in giant pandas using EIA. Although our findings revealed that the excreted steroid metabolites are not necessarily the intended target compounds of those commonly intended with EIAs, they are probably still detected via cross-reactivity of the antibodies. These results are important for future optimization of employed EIAs and will as such contribute to improving the routine reproductive monitoring of giant pandas as part of their conservation efforts.

Data availability

All statistical data are available in the manuscript and supplementary information. The datasets generated and/or analyzed during the current study are available in the Metabolomics Workbench Repository, Study ID: ST003782

References

1. He, K. et al. Effects of roads on giant panda distribution: a mountain range scale evaluation. *Sci. Rep.* **9**, 1110 (2019).
2. Tang, J. et al. Climate change and landscape-use patterns influence recent past distribution of giant pandas. *Proc. R Soc. B: Biol. Sci.* **287**, 20200358 (2020).
3. Kang, D. et al. Feeding habitat characteristics of giant pandas at different scales: A case study in the Wanglang nature reserve. *Glob Ecol. Conserv.* **17**, e000542 (2019).
4. Zeng, Y. et al. Feeding habits and foraging patch selection strategy of the giant panda in the Meigu Dafengding National nature reserve, Sichuan province, China. *Environ. Sci. Pollut Res.* **30**, 49125–49135 (2023).
5. Swaisgood, R. R., Zhou, X., Zhang, G., Lindburg, D. G. & Zhang, H. Application of behavioral knowledge to conservation in the giant panda. *Int. J. Comp. Psychol.* **16**, 65–84 (2003).
6. Lindburg, D. G., Czekala, N. M. & Swaisgood, R. R. Hormonal and behavioral relationships during estrus in the giant panda. *Zoo Biol.* **20**, 537–543 (2001).
7. Zhang, B. et al. Genetic viability and population history of the giant panda, putting an end to the 'evolutionary dead end'? *Mol. Biol. Evol.* **24**, 1801–1810 (2007).
8. Shen, F. et al. Microsatellite variability reveals the necessity for genetic input from wild giant pandas (*Ailuropoda melanoleuca*) into the captive population. *Mol. Ecol.* **18**, 1061–1070 (2009).
9. Huang, Y. et al. Relationship of the Estrogen surge and multiple mates to Cub paternity in the giant panda (*Ailuropoda melanoleuca*): implications for optimal timing of copulation or artificial insemination. *Biol. Reprod.* **87**, 112, 1–7 (2012).
10. Roberts, B. M. et al. Use of urinary 13,14, dihydro-15-keto-prostaglandin F2α (PGFM) concentrations to diagnose pregnancy and predict parturition in the giant panda (*Ailuropoda melanolecua*). *PLoS One.* **13**, e0195599 (2018).
11. Wilson, K. S. et al. Urinary estrogens as a non-invasive biomarker of viable pregnancy in the giant panda (*Ailuropoda melanoleuca*). *Sci. Rep.* **9**, 12772 (2019).
12. Bonney, R. C., Wood, D. J. & Kleimant, D. G. Endocrine correlates of behavioural oestrus in the female giant panda (*Ailuropoda melanoleuca*) and associated hormonal changes in the male. *Journals Reprod. Fertility.* **64**, 209–215 (1982).
13. Wilson, K. S. et al. The novel use of urinary androgens to optimise detection of the fertile window in giant pandas. *Reprod. Fertil.* **3**, 122–132 (2022).
14. Mainka, S. A., Cooper, R. M., Mao, L. & Guanlu, Z. Urinary hormones in two juvenile female giant pandas (*Ailuropoda melanoleuca*). *J. Zoo Wildl. Med.* **21**, 334–341 (1990).
15. Zhang, H., Li, D., Wang, C. & Hull, V. Delayed implantation in giant pandas: the first comprehensive empirical evidence. *Reproduction* **138**, 979–986 (2009).
16. Kersey, D. C. et al. Unique biphasic progestagen profile in parturient and non-parturient giant pandas (*Ailuropoda melanoleuca*) as determined by faecal hormone monitoring. *Reproduction* **140**, 183–193 (2010).
17. Wauters, J. et al. Pregnancy length and health in giant pandas: what can metabolic and urinary endocrine markers unveil? *Theriogenology Wild.* **3**, 100063 (2023).
18. Wauters, J. et al. Evolutionary survival strategies of the female giant panda: optimizing energy resources and expenditure prior to pregnancy by postponing corpus luteum reactivation. *BioRxiv* https://doi.org/10.1101/2022.06.28.497916 (2022).
19. Hodges, J. K. et al. Aspects of the reproductive endocrinology of the female giant panda (*Ailuropoda metanoleaca*) in captivity with special reference to the detection of ovulation and pregnancy. *J. Zool.* **203**, 253–267 (1984).
20. Yamaguchi, N., Dugdale, H. L. & Macdonald, D. W. Female receptivity, embryonic diapause, and superfetation in the European Badger (*Meles meles*): implications for the reproductive tactics of males and females. *Q. Rev. Biol.* **81**, 33–48 (2006).
21. Herrick, J. R. Assisted reproductive technologies for endangered species conservation: developing sophisticated protocols with limited access to animals with unique reproductive mechanisms. *Biol. Reprod.* **100**, 1158–1170 (2019).
22. Willis, E. L., Kersey, D. C., Durrant, B. S. & Kouba, A. J. The acute phase protein ceruloplasmin as a non-invasive marker of pseudopregnancy, pregnancy, and pregnancy loss in the giant panda. *PLoS One.* **6**, e21159 (2011).
23. Wauters, J. et al. Urinary specific gravity as an alternative for the normalisation of endocrine metabolite concentrations in giant panda (*Ailuropoda melanoleuca*) reproductive monitoring. *PLoS One.* **13**, e0201420 (2018).
24. De Paepe, E. et al. A validated multi-matrix platform for metabolomic fingerprinting of human urine, feces and plasma using ultra-high performance liquid-chromatography coupled to hybrid orbitrap high-resolution mass spectrometry. *Anal. Chim. Acta.* **1033**, 108–118 (2018).
25. Cools, T. et al. Development and validation of a versatile non-invasive urinary steroidomics method for wildlife biomonitoring. *Talanta* **273**, 125924 (2024).
26. Vangeenderhuysen, P. et al. Dual UHPLC-HRMS metabolomics and lipidomics and automated data processing workflow for comprehensive High-Throughput gut phenotyping. *Anal. Chem.* **95**, 8461–8468 (2023).
27. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple. *J. R Stat. Soc. Ser. B.* **57**, 289–300 (1995).
28. Dührkop, K. et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods.* **16**, 299–302 (2019).
29. Aron, A. T. et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **15**, 1954–1991 (2020).
30. Schymanski, E. L. et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
31. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R Stat. Soc. Ser. B.* **67**, 301–320 (2005).
32. Cho, S., Kim, H., Oh, S., Kim, K. & Park, T. Elastic-net regularization approaches for genome-wide association studies of rheumatoid arthritis. *BMC Proc* **3**, (2009).
33. Maleki, F., Muthukrishnan, N., Ovens, K., Reinhold, C. & Forghani, R. Machine learning algorithm validation: from essentials to advanced applications and implications for regulatory certification and deployment. *Neuroimaging Clin. N Am.* **30**, 433–445 (2020).
34. Krupp, D., Doberstein, N., Shi, L. & Remer, T. Hippuric acid in 24-hour urine collections is a potential biomarker for fruit and vegetable consumption in healthy children and adolescents. *J. Nutr.* **142**, 1314–1320 (2012).
35. Vane, J. R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nat. New. Biol.* **231**, 232–235 (1971).
36. Smith, M. J. H., Ford-Hutchinson, A. W., Walker, J. R. & Slack, J. A. Aspirin, salicylate and prostaglandins. *Agents Actions.* **9**, 483–487 (1979).
37. Lim, H. et al. Multiple female reproductive failures in cyclooxygenase 2-Deficient mice. *Cell* **91**, 197–208 (1997).

38. Cao, M. et al. Assessing urinary metabolomics in giant pandas using chromatography/mass spectrometry: Pregnancy-Related changes in the metabolome. *Front. Endocrinol.* **11**, 215 (2020).
39. Wang, M. et al. Normal pregnancy-induced amino acid metabolic stress in a longitudinal cohort of pregnant women: novel insights generated from UPLC-QTOFMS-based urine metabolomic study. *Metabolomics* 12, (2016).
40. Fukuwatari, T. et al. Changes in the urinary excretion of the metabolites of the Tryptophan-Niacin pathway during pregnancy in Japanese women and rats. *J. Nutr. Sci. Vitaminol.* **50**, 392–398 (2004).
41. Foster, A. C., Vezzani, A., French, E. D. & Schwarz, R. Kynurenic acid blocks neurotoxicity and seizures induced in rats by the related brain metabolite quinolinic acid. *Neurosci. Lett.* **48**, 273–278 (1984).
42. Mándi, Y. & Vécsei, L. The kynurenine system and immunoregulation. *J. Neural Transm.* **119**, 197–209 (2012).
43. Suzuki, Y. et al. Identification of antioxidants produced by *Lactobacillus plantarum*. *Biosci. Biotechnol. Biochem.* **77**, 1299–1302 (2013).
44. Oliver, A. et al. Cervicovaginal Microbiome composition is associated with metabolic profiles in healthy pregnancy. *mBio* **11**, 10–1128 (2020).
45. Laursen, M. F. et al. *Bifidobacterium* species associated with breastfeeding produce aromatic lactic acids in the infant gut. *Nat. Microbiol.* **6**, 1367–1382 (2021).
46. Dambrova, M. et al. Acylcarnitines: nomenclature, biomarkers, therapeutic potential, drug targets, and clinical trials. *Pharmacol. Rev.* **74**, 506–551 (2022).
47. Houten, S. M. & Wanders, R. J. A. A general introduction to the biochemistry of mitochondrial fatty acid β-oxidation. *J. Inherit. Metab. Dis.* **33**, 469–477 (2010).
48. Xu, G. et al. Liver and muscle contribute differently to the plasma acylcarnitine pool during fasting and exercise in humans. *J. Clin. Endocrinol. Metab.* **101**, 5044–5052 (2016).
49. Manta-Vogli, P. D., Schulpis, K. H., Dotsikas, Y. & Loukas, Y. L. The significant role of carnitine and fatty acids during pregnancy, lactation and perinatal period. Nutritional support in specific groups of pregnant women. *Clin. Nutr.* **39**, 2337–2346 (2020).
50. Zeng, X. et al. The combined use of Medium- and Short-Chain fatty acids improves the pregnancy outcomes of sows by enhancing ovarian steroidogenesis and endometrial receptivity. *Nutrients* **14**, 4405 (2022).
51. Ye, Q. et al. Maternal short and medium chain fatty acids supply during early pregnancy improves embryo survival through enhancing progesterone synthesis in rats. *J. Nutr. Biochem.* **69**, 98–107 (2019).
52. Bobiński, R. & Mikulska, M. The Ins and outs of maternal-fetal fatty acid metabolism. *Acta Biochim. Pol.* **62**, 499–507 (2015).
53. Kersey, D. C. et al. Adrenal hormone concentrations in the female giant panda (*Ailuropoda melanoleuca*), as measured by fecal glucocorticoids, are associated with reproductive activity. *Biol Reprod* **81**, (2009).
54. Kersey, D. C. et al. Rising fecal glucocorticoid concentrations track reproductive activity in the female giant panda (*Ailuropoda melanoleuca*). *Gen. Comp. Endocrinol.* **173**, 364–370 (2011).
55. Jenkin, G. & Young, I. R. Mechanisms responsible for parturition; the use of experimental models. *Anim. Reprod. Sci.* **82–83**, 567–581 (2004).
56. Liggins, G. C. The role of cortisol in Preparing the fetus for birth**. *Reprod. Fertil. Dev.* **6**, 141–150 (1994).
57. Whittle, W. L. et al. Glucocorticoid regulation of human and ovine parturition: the relationship between fetal Hypothalamic-Pituitary-Adrenal axis activation and intrauterine prostaglandin production. *Biol. Reprod.* **64**, 1019–1032 (2001).
58. Norris, D. O. & Carr, J. A. *Vertebrate Endocrinology* (Academic, 2021).
59. Cloke, B. & Christian, M. The role of androgens and the androgen receptor in cycling endometrium. *Mol. Cell. Endocrinol.* **358**, 166–175 (2012).
60. Diao, H. L. et al. Effects of androgen on embryo implantation in the mouse delayed-implantation model. *Fertil. Steril.* **90**, 1376–1383 (2008).

Acknowledgements

Author contributions
TC has performed the formal analysis and investigation of the data. JW was responsible for the conceptualization of the study, while LV provided resources for the experiments. Both LV and JW were supervising the study. TC and KW visualized the data, and TC was also responsible for the original draft of the manuscript together with JW and LV. Samples and approval for the study were provided by DL, CV, BM, AL, JK, MH, MB, AO, SG, YZ, and RL. All authors have read and approved the final manuscript.

Funding

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-19067-7.

Correspondence and requests for materials should be addressed to L.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.