Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Exploring the molecular regulatory network of non-pathological fetal overgrowth phenotype in Holstein cattle derived from in vitro embryo production

Leiyu Qu
Yanbian University
Zhengyu Hang
China Agricultural University
Yao Chang
China Agricultural University
Lei Wang
China Agricultural University
Ruina Ma
Beijing Jiaotong University
Shanjiang Zhao
Institute of Animal Science, Chinese Academy of Agricultural Sciences
Qingshan Gao
Yanbian University
Qing Xu
Beijing Jiaotong University
Yachun Wang
wangyachun@cau.edu.cn
China Agricultural University

Research Article

Keywords: In vitro embryo production, Holstein cattle, multi-omics analysis, fetal overgrowth

Posted Date: August 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10534886/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on October 6th, 2026. See the published version at https://doi.org/10.1186/s12864-026-13408-0.

Exploring the molecular regulatory network of non-pathological fetal overgrowth phenotype in Holstein cattle derived from *in vitro* embryo production

Leiyu Qu[1], Zhengyu Hang[2], Yao Chang[2], Lei Wang[2], Ruina Ma[3], ShanJiang Zhao[4], Qingshan Gao[1,*], Qing Xu[3,*], Yachun Wang[2,*]

[1]College of Agriculture, Yanbian University, Yanji, 133000, China

[2]State Key Laboratory of Animal Biotech Breeding, National Engineering Laboratory for Animal Breeding, Laboratory of Animal Genetics, Breeding, and Reproduction, MARA, College of Animal Science and Technology, China Agricultural University, Beijing, 100193, China

[3]College of Life Sciences and Bio-Engineering, Beijing Jiaotong University, Beijing, 100044, China

[4]State Key Laboratory of Animal Genetics, Breeding and Reproduction of Ministry of Agriculture and Rural Affairs, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

*Corresponding author

wangyachun@cau.edu.cn

E-mail addresses:

LQ: gubang2019@163.com

ZH: zhengyu.hang@outlook.com

YC: changyaocy@qq.com

LW:wledu2016@163.com

NM: 1526099015@qq.com

JZ: zhaoshanjiang@caas.cn

GQ:qsgao@ybu.edu.cn

27 QX:qingxu@bjtu.edu.cn

28 YW: wangyachun@cau.edu.cn

29

Abstract

Background

Calves derived from *in vitro* embryo production (IVP) often have higher birth weights than those produced by artificial insemination or *in vitro* embryo transfer, and are frequently accompanied by malformations, mortality, or dystocia in dams. However, a subset of calves exhibits only increased body weight without adverse health outcomes. This study focused on the non-pathological fetal overgrowth phenotype, aiming to refine the understanding of abnormal birth weight in IVP calves and provide new perspectives and a theoretical basis for the underlying regulatory mechanisms.

Results

In this study, whole-transcriptome, metabolomic, physiological, biochemical, and hormonal analyses were integrated to compare nine IVP calves with the non-pathological fetal overgrowth phenotype and nine IVP calves with normal birth weight. The results demonstrated that increased birth weight in IVP calves was accompanied by prolonged gestation length in recipient dams, and multiple physiological, biochemical, and hormonal parameters were significantly altered. Untargeted metabolomics identified 3,259 metabolites, among which 132 differential metabolites were screened, including 82 significantly up-regulated and 50 significantly down-regulated metabolites. Two core regulatory axes were identified in the ceRNA network: circ36279-bta-miR-431 (targeting SPON2, TNN, INSIG1, and AEN) and circ35151-bta-miR-154b (targeting SHD, TBX2, and IL20RA). Integrative analysis revealed 5 co-enriched pathways, with arginine and proline metabolism as the core pathway, mediating multiple downstream branches including polyamine synthesis, ornithine metabolism, creatine synthesis, and proline metabolism through key enzymes such as AGMAT, ARG1, and CKM, collectively constituting a complex regulatory network.

Conclusions

Through multi-omics integrative analysis, this study clarified that IVP-associated non-pathological fetal overgrowth is accompanied by systemic physiological and biochemical changes and identified two ceRNA regulatory networks and arginine and proline metabolism as a core pathway. These findings refine the molecular understanding of this phenotype and provide new perspectives for precise intervention in growth abnormalities of IVP calves.

Keywords *In vitro* embryo production, Holstein cattle, multi-omics analysis, fetal overgrowth

Background

In vitro embryo production technology can rapidly expand the core cow population and shorten the generation interval [1]. According to data from the International Embryo Technology Society, the total number of embryos collected or produced globally in 2023 was 2,411,329, representing a 13.2%

increase compared with 2022. Among these, the number of bovine *in vitro* production (IVP) embryos was 1,876,591, an increase of 15.8% from 2022, while the number of bovine in vivo-derived (IVD) embryos was 371,138, a decrease of 6.3% from 2022. The number of bovine IVP embryos currently exceeds that of IVD embryos by more than fivefold. Based on these figures, IVP bovine embryos accounted for approximately 83.5% of all bovine embryos (2,247,729) in 2023[2].

However, with the large-scale application of IVP technology in dairy farms, abnormalities such as offspring malformations, weak calves, and fetal overgrowth have become increasingly prominent. Large Offspring Syndrome (LOS) is a congenital overgrowth disorder characterized by excessive body size, abdominal wall defects, organomegaly, and difficulty standing and suckling at birth [3]. Overgrown fetuses are at increased risk of dystocia, which can lead to maternal injury or even death, resulting in economic losses to dairy farms. The epigenetic features of LOS include multilocus imprinting loss, global dysregulation of small RNAs and long RNAs, altered DNA methylation, and chromosomal structural abnormalities. Furthermore, ART-induced LOS has been characterized as an imprinting loss syndrome.

Notably, in production practice, a subset of IVP calves exhibits significantly increased birth weight but subsequent normal growth and development without obvious pathological manifestations, representing a non-pathological fetal overgrowth phenotype. Currently, the molecular regulatory mechanisms underlying this phenotype remain unclear. The birth weight of IVP Holstein calves in the study by Eliza M. Murphy et al. was 39.0 ± 1.4 kg [4].Fetal overgrowth is closely associated with alterations in multiple physiological, biochemical, and hormonal parameters. A study by María B. Rabaglino identified 15 epigenetic and five transcriptomic biomarkers that can predict embryo origin using blood epigenomic data[5].

Integrative analysis of the whole transcriptome and metabolome enables systematic dissection of the molecular mechanisms underlying complex phenotypes at both the RNA regulatory and metabolite levels [6]. Whole-transcriptome sequencing allows systematic identification of multiple transcript species, including mRNAs, long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs), and facilitates the construction of competing endogenous RNA (ceRNA) regulatory networks to uncover key molecular events at the post-transcriptional level. Meanwhile, metabolomics, as the "terminal output" of functional phenotypes, precisely captures the end products of biochemical reactions in vivo. The combination of these two approaches provides a feasible technical pathway for in-depth investigation of the intrinsic molecular regulatory mechanisms underlying the aforementioned non-pathological overgrowth.

This study initially performed descriptive statistical analyses of birth weights in Chinese Holstein calves derived from large-scale commercial in vitro embryo production (IVP) farms to select suitable experimental subjects. Subsequently, we integrated physiological and biochemical indices, hormonal assays, metabolomic profiling, and whole-transcriptome sequencing to construct a competing endogenous RNA (ceRNA) regulatory network. Joint analysis of transcriptomic and metabolomic data identified differential pathways predominantly involving arginine and tryptophan metabolism,

thereby offering insights into the potential molecular mechanisms underlying abnormal birth weight in IVP calves. Collectively, these findings refine our molecular-level understanding of IVP-associated birth weight anomalies and provide both empirical data and theoretical support for optimizing IVP production protocols, refining diagnostic criteria for large calf syndrome, and developing targeted intervention strategies.

Methods

Animal Selection and Sample Collection

This study was conducted from June to August 2025 at a large-scale dairy farm with approximately 20,000 cattle in Jilin Province, China.All experimental procedures involving animals were reviewed and approved by the Animal Ethics Committee of China Agricultural University (approval number: A0070505202-1-01). All procedures were conducted in accordance with the Guidelines for the Care and Use of Laboratory Animals to ensure ethical compliance.

For animal enrollment, calves derived from *in vitro* embryos with birth weight exceeding two SD above the population mean and in good health were primarily selected as the experimental BIG group, consisting of a total of nine female calves with a mean birth weight of 56.28 ± 2.33 kg. While healthy nine female calves with birth weights near the population mean (38.39 ± 1.27 kg) were selected as the control group (HCG). All calves in this population were derived from *in vitro* embryo production. In this study, the non-pathological fetal overgrowth referred to in the BIG group was defined based on good postnatal condition and sustained good health beyond sixty days of age in the calves.

All donors, recipients, and sires were purebred Holstein cattle, clinically healthy, with no reproductive, metabolic, or infectious disorders, and complete pedigree records traceable for at least three generations. To minimize technical variation, the entire IVP procedures, including ovum pick-up (OPU), *in vitro* maturation (IVM), *in vitro* fertilization (IVF), embryo culture, and embryo transfer (ET), was performed by a single professional technical team.

After birth, each calf was weighed immediately, and health status, calving assistance score, and birth time were recorded. Blood samples were collected without anesthesia, in compliance with current animal welfare and research laws and regulations in China (Approval code: AW70505202-1-01). 10 mL of jugular venous blood was collected into anticoagulant tubes. Blood samples were centrifuged at 3,500 rpm for 15 min. Plasma was transferred into 2-mL cryovials and stored for subsequent analyses. The buffy coat layer was collected, and 1 mL of buffy coat was transferred into a 2-mL cryovial containing 1 mL of TRIzol reagent (Thermo Fisher Scientific Inc., US). The samples were thoroughly mixed using a vortex mixer and immediately stored in liquid nitrogen at -80°C. All samples were subsequently shipped on dry ice.

Physiological, Biochemical and Hormonal Measurements

Twenty blood samples in total were measured for eighteen physiological and biochemical parameters. Superoxide dismutase (SOD), catalase (CAT), total antioxidant capacity (T-AOC), and glutathione peroxidase (GSH-Px) were measured using biochemical assay kits and an L-3180 semi-automatic biochemical analyzer (Shanghai Kehua Bio-Engineering co., Ltd., Shanghai, China). Albumin (ALB), alanine aminotransferase (ALT), total cholesterol (TC), triglyceride (TG), high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), creatinine (CRE), immunoglobulin A (IgA), and immunoglobulin G (IgG) were measured using a URIT CA-431A automatic biochemical analyzer (URIT Medical Electronics Co., Guilin, China). Growth hormone (GH) and insulin-like growth factor 1 (IGF-1) were quantified by enzyme-linked immunosorbent assay using an ST-360 microplate reader (Shanghai Kehua Bio-Engineering co., Ltd., Shanghai, China). Additional measurements included non-esterified fatty acids (NEFA), β-hydroxybutyrate (β-BHBA), interleukin-6 (IL-6), interleukin-2 (IL-2), reactive oxygen species (ROS), neutrophil elastase (NE), brain-derived neurotrophic factor (BDNF). All analyses were performed by Beijing Jin Hai Ke Yu Biological Technology Development Co., Ltd. (Beijing, China), and each assay was conducted according to the manufacturer's instructions.

Metabolite Extraction

For metabolite extraction, 100 μL of plasma was added to a 1.5-mL centrifuge tube containing 400 μL of extraction solution consisting of acetonitrile and methanol at a ratio of 1:1 (v/v). The extraction solution contained four internal standards, including 0.02 mg/mL L-2-chlorophenylalanine. Samples were vortexed for 30 s and subjected to low temperature ultrasonication for 30 min at 5°C and 40 kHz. The samples were then incubated at -20°C for 30 min to precipitate proteins, followed by centrifugation at 13,000 g/min for 15 min at 4°C. The supernatant was collected and dried under nitrogen gas.

The dried extracts were reconstituted in 120 μL of acetonitrile: water solution at a ratio of 1:1 (v/v) and extracted by low temperature ultrasonication for 5 min at 5°C and 40 kHz. The samples were then centrifuged at 13,000 g/min for 10 min at 4°C. The resulting supernatant was transferred to sample vials for LC-MS/MS analysis.

Differential metabolites identified between the BIG and HCG were mapped to biochemical pathways using the KEGG database. Pathway enrichment and topological analyses were performed to identify significantly altered metabolic processes.

RNA Extraction, Library Preparation and Sequencing

Total RNA was extracted from buffy coat samples using TRIzol reagent. RNA quality was assessed using a 5300 Bioanalyzer (Agilent Technologies, USA), and RNA concentration was quantified using

an ND-2000 spectrophotometer (NanoDrop Technologies, USA). Only high-quality RNA samples meeting the following criteria were used for library construction: OD260/280 = 1.8-2.2, OD260/230 ≥ 2.0, RNA integrity number (RIN) ≥ 6.5, 28S:18S ratio ≥ 1.0, and total RNA amount >1 μg.

Small RNA libraries were constructed using the QIAseq miRNA Library Kit (Qiagen, Germany). Ribosomal RNA was removed using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus Kit (Illumina, San Diego, CA, USA), and mRNA, lncRNA, and circRNA libraries were subsequently constructed for long RNA sequencing. All prepared libraries were sequenced on the Illumina NovaSeq 6000 platform (Illumina, USA).

For functional enrichment analysis of lncRNAs, mRNAs located within 100 kb upstream or downstream of each lncRNA were selected as potential cis-acting target genes[7]. Differentially expressed lncRNAs and mRNAs were identified using the criteria $|\log_2(\text{fold change})| > 1$ and $P < 0.05$. mRNAs located within 100 kb of differentially expressed lncRNAs were considered potential lncRNA target genes.

Bioinformatic Analysis and RNA Identification

Clean reads were aligned to the reference genome using HISAT2 v2.1.0 [8]. The expression levels of mRNAs were quantified using featureCounts and normalized as transcripts per million (TPM) [9]. For lncRNA identification, transcripts were reconstructed using StringTie v2.1.1[10]. A stringent filtering pipeline was applied to identify novel lncRNAs. Transcripts with a length >200 bp and at least two exons were first retained. Coding potential was then evaluated using CPC0.9-r2 (http://cpc2.cbi.pku.edu.cn) and CNCI2.0 (https://github.com/www-bioinfo-org/CNCI#install-cnci), and only transcripts predicted to lack coding potential by both tools were defined as lncRNAs.

For circRNA identification, reads that could not be aligned to the reference genome were further analyzed using back-splice alignment algorithms, including find_circ and CIRI2 to identify circular junctions [11, 12]. circRNA expression levels were quantified as reads per million mapped reads (RPM) based on the number of back-splicing junction (BSJ) reads.

For small RNA sequencing data, clean tags were first filtered by length, typically retaining reads of 18-30 nucleotides. The processed tags were aligned against the miRBase v22.0 database to identify known miRNAs[13]. Unannotated reads were then analyzed using miRDeep2 to predict novel miRNAs based on the characteristic hairpin structures of miRNA precursors[14].

Differential Expression and Functional Enrichment Analysis

Principal component analysis (PCA) was performed separately for the normalized mRNA, lncRNA, circRNA, and miRNA expression datasets to assess overall expression patterns and sample-level variation between the HCG and BIG groups. PCA was conducted using the prcomp function in R. The first two principal components were used for visualization.

Differential expression analysis of mRNAs, lncRNAs, circRNAs, and miRNAs was performed using R v4.5.2 with the DESeq2 package [15]. Raw read count data obtained from expression quantification were used as input for differential expression analysis. Differentially expressed genes (DEGs), differentially expressed lncRNAs (DELs), differentially expressed circRNAs (DECs), and differentially expressed miRNAs (DEMs) were identified using the criteria $|\log_2(\text{fold change})| > 1$ and $P < 0.05$.

To explore the biological functions of the identified DEGs, Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analyses were conducted using KOBAS [16]. GO terms and KEGG pathways with $P < 0.05$ were considered significantly enriched.

A protein-protein interaction (PPI) network was constructed for proteins encoded by the differentially expressed genes using the STRING database (version 12.0; https://string-db.org/), with Bos taurus selected as the reference organism. Interactions with a combined confidence score ≥ 0.7 were retained, and disconnected nodes were excluded. The interaction data were imported into Cytoscape (version 3.9.1) for network visualization. Node connectivity was calculated using the Degree algorithm in the CytoHubba plugin, and genes with the highest Degree values were identified as candidate hub genes. Node color and size were mapped to Degree values, with darker and larger nodes representing higher connectivity.

miRNA Target Gene Prediction

Target genes of differentially expressed miRNAs were predicted using miRanda v3.3a [17] and RNAhybrid v2.1.2 [18]. To improve the reliability of the predicted miRNA-target interactions, only the overlapping results obtained from both algorithms were retained for further analysis [12]. For miRanda, the binding energy threshold was set to ≤-20 kcal/mol, with $P < 0.05$. For RNAhybrid, a minimum free energy threshold of ≤-20 kcal/mol was applied to ensure stable miRNA-target duplexes.

Weighted Gene Co-expression Network Analysis

Weighted Gene Co-expression Network Analysis (WGCNA) was performed separately for the mRNA and lncRNA expression profiles using the WGCNA package (version 1.69) in R [19]. Normalized expression values were transformed as $\log_2(x + 1)$, and the top 50% of transcripts ranked by median absolute deviation (MAD) were retained for network construction.

Signed co-expression networks were constructed using soft-thresholding powers selected according to the scale-free topology fit index and mean connectivity. Adjacency matrices were converted into topological overlap matrices, followed by hierarchical clustering. Co-expression modules were identified using the dynamic tree-cutting algorithm with a minimum module size of 30, and modules with similar eigengene profiles were merged using a mergeCutHeight of 0.25. Module eigengenes (MEs) were calculated to represent the expression profile of each module.

Competing Endogenous RNA (ceRNA) Network Construction

A circRNA-miRNA-mRNA regulatory network was constructed to investigate potential ceRNA interactions. miRNA binding sites in differentially expressed circRNAs were predicted using miRanda (http://mirtoolsgallery.tech/mirtoolsgallery/node/1055), and interactions with a binding energy ≤-20 kcal/mol were retained. Target mRNAs of miRNAs were predicted using both miRanda and RNAhybrid [18]. Only overlapping target genes identified by both algorithms were included in the network. The circRNA-miRNA-mRNA regulatory network was visualized using Cytoscape v3.2.0 (https://cytoscape.org/release_notes_3_2_0.html).

Statistical Analysis

All statistical analyses were performed using SAS software (version 9.4; SAS Institute Inc., Cary, NC, USA). Figures were generated using R software (R Core Team, Vienna, Austria), and tables were prepared using Microsoft Excel (Microsoft Corp., Redmond, WA, USA). Data were analyzed using the GLM procedure in SAS. Differences were considered statistically significant at $P < 0.05$ and highly significant at $P < 0.01$.

Results

Intergroup differences in calf birth weight and related phenotypic indicators of recipient dams

In this study, female calves were selected as the research subjects, and a descriptive statistical analysis of birth weight was performed on 4,482 IVP female calves, with a population mean birth weight of 39.83 ± 5.94 kg. Based on the calculation of standard deviations, the thresholds for one SD, two SD, and three SD above the mean were 45.76 kg, 51.69 kg, and 57.62 kg, respectively. Violin plots combined with significance tests were used to compare the differences in four phenotypic indicators between the HCG and BIG groups: calf birth weight, gestation length of recipient dams, postpartum body weight of recipient dams, and postpartum body height of recipient dams. Statistical results showed (Fig. 1) that calf birth weight in the BIG group 56.28±2.33kgwas extremely significantly higher than that in the HCG control group 38.39±1.27kg (P<0.01) . The gestation length of recipient dams in the BIG group 281.67±9.66d was significantly longer than that in the HCG control group 272.78±3.49d (P<0.05) , indicating a marked prolongation of the gestational period in the experimental group. No statistically significant difference was observed in postpartum body weight between the HCG and BIG recipient dam groups (664.33 ± 40.26 kg vs. 634.78 ± 53.22 kg; $P >$ 0.05).Postpartum body height of recipient dams in the BIG group 152.89±2.85cm was significantly higher than that in the HCG control group 149.33±4.09cm (P<0.05).

Differential analysis of blood physiological, biochemical, and neuroendocrine hormone levels in calves

In this study, violin plots were used to visualize various physiological, biochemical, and hormonal parameters in blood from nine HCG control calves and nine BIG experimental calves, and intergroup differences were tested for significance (Fig. 2). Results of routine serum biochemical and liver and kidney function indicators showed that the liver function indicators ALB and ALT were significantly higher in the BIG group than in the HCG group ($P<0.05$) ; the kidney function marker CRE was significantly higher in the HCG group than in the BIG group ($P<0.05$) Lipid metabolism and energy homeostasis indicators showed that TC, HDL-C, and LDL-C levels were significantly increased in the BIG group ($P<0.05$) ; NE, which reflects fat mobilization, was extremely significantly higher in the BIG group than in the HCG group ($P<0.01$) , while the ketone body β-BHBA was extremely significantly lower in the BIG group than in the HCG group ($P<0.01$). Oxidative stress and antioxidant defense system results showed that ROS levels were significantly decreased in the BIG group ($P<0.05$) ; the antioxidant enzymes SOD, CAT, and GSH-Px, as well as total antioxidant capacity (T-AOC), were all extremely significantly lower in the BIG group than in the HCG group ($P<0.01$) . Immune response and inflammatory cytokine results showed that the humoral immunity indicators IgG and IgA were significantly lower in the BIG group than in the HCG group ($P<0.01$) ; the pro-inflammatory cytokine IL-6 and the immune regulatory factor IL-2 were highly significantly higher in the BIG group than in the HCG group ($P<0.01$) . Neuroendocrine regulation and neuroplasticity factor results showed significant intergroup differences in growth axis-related hormones, with GH, IGF-1, and NE being extremely significantly higher in the BIG group than in the HCG group ($P<0.01$).

Blood metabolomic profiling and differential metabolite analysis

The classification statistics of metabolites identified in blood samples showed (Fig. 3A) that lipids and lipid-like molecules accounted for the highest proportion (27.27%), followed by organic acids and their derivatives (26.31%). The Venn diagram illustrates the distribution of detected metabolites between the HCG and BIG groups (Fig. 3B) . A total of 3,019 metabolites were shared by both groups. The HCG group contained 190 unique metabolites, while the BIG group contained 95 unique metabolites. PCA of untargeted metabolomics showed (Fig. 3C) that the first principal component (PC1) explained 26.38% of the total variance, and the second principal component (PC2) explained 14.07%, with the two groups showing clear separation. Volcano plot analysis of untargeted metabolomics showed (Fig. 3D) that a total of 3,259 metabolites were identified in the two groups of Holstein calves, and 132 differential metabolites were screened, including 82 significantly up-regulated and 50 significantly down-regulated metabolites. The up-regulated markers in the BIG group were mainly amino acids and organic acids (Ser-Thr-Ile, 3-phosphopropionic acid, etc.), while the down-regulated substances were mostly various phospholipids and lipids. The KEGG pathway enrichment bubble plot results showed (Fig. 3E) that alanine, aspartate and glutamate metabolism, bile secretion, D-amino acid metabolism, and arginine and proline metabolism showed the highest enrichment intensity and significance. Amino acid metabolism, lipid digestion, insulin resistance, and neuroactive ligand-receptor interaction pathways together constituted the differential metabolic regulatory network, with arginine and proline metabolism serving as the core pathway regulating overgrowth and development in IVP calves.

Blood mRNA transcriptome sequencing analysis

Based on the gene expression data from mRNA transcriptome sequencing, both sample hierarchical clustering and PCA results showed (Fig. 4A) a clear grouping trend between the two groups in gene expression space. DEG analysis identified a total of 1,100 DEGs, including 567 up-regulated genes and 533 down-regulated genes; *DPH2*, *NFKBID*, and *P2RY11* showed particularly significant expression differences between groups (Fig. 4B). The protein-protein interaction (PPI) network results showed (Fig. 4C) that the top five core genes, ranked by node connectivity (degree), were *ALB*, *IL1B*, *IL10*, *FGF2*, and *HIF1A*. Weighted gene co-expression network analysis (WGCNA) identified a total of seven modules (Fig. 4D), with the turquoise module being the most correlated module, containing 1,307 genes. GO functional annotation and KEGG pathway enrichment analysis of all genes within this module showed (Fig. 4E) that biological processes were mainly enriched in cell adhesion, erythrocyte development, oxidative metabolism, oxygen transport, and cell differentiation; cellular components were mainly localized to the extracellular region, cell membrane, and cytoskeleton; molecular functions were significantly enriched in ion binding, oxidase activity, and transcriptional regulatory activity. KEGG pathway enrichment analysis indicated that the module genes were significantly involved in extracellular matrix (ECM)-receptor interaction, focal adhesion, PI3K-Akt signaling pathway, cytokine-cytokine receptor interaction, and insulin secretion.

Blood lncRNA transcriptome sequencing analysis

In the lncRNA transcriptome sequencing of this study, principal component analysis results showed (Fig. 5A) that the lncRNA expression profiles of the two groups exhibited a clear inter-group separation trend, with PCA1 and PCA2 explaining 14.78% and 10.43% of the total variation, respectively. Volcano plot results of differential lncRNAs showed (Fig. 5B) that a total of 562 differentially expressed lncRNAs were identified, including 239 up-regulated lncRNAs and 323 down-regulated lncRNAs. The soft-threshold fitting index results of lncRNA WGCNA showed (Fig. 5C) that when the soft threshold reached two, the Scale Free R^2 had already exceeded 0.8, and as the threshold continued to increase, the fitting index remained stable above 0.9. WGCNA analysis identified a total of 4 modules significantly associated with grouping (Fig.5D), among which the turquoise module showed the highest correlation and was the most significant module, containing 865 lncRNAs. Intersection analysis between lncRNAs within the WGCNA significant module and global differentially expressed lncRNAs identified 380 cores differentially expressed lncRNAs (Fig. 5E). Further KEGG pathway enrichment analysis was performed on 841 adjacent genes within 100 kb upstream and downstream of the 380 core differentially expressed lncRNAs, and the results showed that these genes were mainly enriched in cornified envelope formation, metabolic pathways, GABAergic synapse, collecting duct acid secretion, efferocytosis, gastric acid secretion, leukocyte transendothelial migration, estrogen signaling pathway, galactose metabolism, and fructose and mannose metabolism (Fig. 5F).

Blood miRNA and circRNA transcriptome and ceRNA regulatory network analysis

Blood samples from nine HCG controls and nine BIG experimental calves were subjected to miRNA and circRNA transcriptome sequencing to characterize the regulatory features. PCA of miRNAs showed that samples from the HCG and BIG groups formed distinct group clustering separation, with PCA1 and PCA2 explaining 26.18% and 16.47% of the total variation, respectively, indicating significant differences in the overall miRNA expression profiles between the two groups (Fig. 6A). Volcano plot results (Fig. 6B) identified a total of 34 differentially expressed miRNAs, including 21 up-regulated and 13 down-regulated miRNAs. PCA at the circRNA level similarly showed clear inter-group discrimination (Fig. 6C), with PCA1 and PCA2 contributing 29.43% and 7.12% of the total variance, respectively. circRNA volcano plot results (Fig. 6D) identified a total of 471 significantly differentially expressed circRNAs, including 300 up-regulated and 171 down-regulated circRNAs. To explore the transcriptional regulatory mechanisms of non-coding RNAs, a circRNA-miRNA-mRNA ceRNA regulatory network was constructed based on targeting binding relationships (Fig. 6E). Two core regulatory axes were identified in the network: circ36279-bta-miR-431 and circ35151-bta-miR-154b. Among these, bta-miR-431 could target and bind multiple mRNAs including *SPON2*, *TNN*, *INSIG1*, and *AEN*, while bta-miR-154b could target and regulate the downstream target mRNAs *SHD*, *TBX2*, and *IL20RA*.

Integrative analysis of WGCNA modules and metabolomics, and gene-metabolite co-expression network

To dissect the synergistic regulatory relationships between the transcriptome and metabolome, we first performed an intersection analysis between the KEGG pathways enriched by genes in the turquoise module from the transcriptomic WGCNA and the KEGG enrichment pathways from metabolomics. The results showed (Fig. 6F) that a total of 5 overlapping pathways were identified (Table 1): arginine and proline metabolism, tryptophan metabolism, neuroactive ligand-receptor interaction, protein digestion and absorption, and mineral absorption. On this basis, Pearson correlation analysis was further performed on the differential genes and differential metabolites contained in the above intersecting pathways, and the association patterns were presented as a heatmap. The results showed (Fig. 7) a significant bidirectional association between differential genes and differential metabolites. Specifically, genes such as *ADRB2*, *P4HA3, ARG1*, *GABRR1*, and *TPH2* showed significant positive correlations with metabolites including creatinine, L-cysteine, L-alanine, L-proline, and adenosine 5'-monophosphate (with the highest correlation coefficient $r = 0.79$); while genes such as *DRD5*, *INSL3*, *DDC*, and *GRIN2C* showed significant negative correlations with the above metabolites (with the lowest correlation coefficient $r = -0.84$). Conversely, metabolites such as D-octopamine, 3-indoleacetaldehyde, L-glutamic-γ-semialdehyde, and L-aspartic acid showed significant positive correlations with genes including *DRD5*, *INSL3*, *DDC*, and *GRIN2C* (r up to 0.78), while showing significant negative correlations with genes such as *ADRB2* and *TPH2* (r as low as -0.71). These results indicate the existence of complex bidirectional regulatory relationships between the transcriptome and metabolome in the related pathways.

Regulatory changes in the arginine and proline metabolism pathway

This study systematically analyzed the regulatory changes in the metabolic network centered on arginine and proline metabolism, along with the expression changes of key enzymes and metabolites. The pathway structure and regulatory patterns are shown in Fig. 8. Arginine and proline metabolism, as a core node, primarily flows through multiple downstream branches, including polyamine synthesis, ornithine metabolism, creatine synthesis, and proline metabolism, via key enzymes such as *AGMAT*, *ARG1*, and *CKM*, forming a complex regulatory network. The expression of *AGMAT* and ARG1 was significantly downregulated, suggesting that the conversion of arginine to agmatine and ornithine was attenuated, and polyamine synthesis was inhibited. In creatine metabolism, the expression of *CKM* showed no significant change, but the abundance of creatinine was significantly elevated, indicating enhanced metabolic flux through this pathway. Ornithine, as a hub, is converted to *P5C* via *OAT* or to pyruvate via *HOGA1*; among these, *HOGA1* expression was downregulated, but pyruvate abundance was upregulated, suggesting activation of glucose metabolism. In proline metabolism, although the expression of key enzymes *P4HA3* and *MAOB* was downregulated, L-proline abundance was significantly elevated, accompanied by a slight upregulation of *P5C*, suggesting that proline accumulation may result from enhanced synthesis or inhibited degradation.

Discussion

In IVP calves, birth weight is the most intuitive indicator for assessing fetal overgrowth. From a statistical perspective, two standard deviations above the population mean birth weight (> mean + 2 SD) can be used to define such overgrowth, and this approach has been adopted by some studies to define fetal overgrowth [20]. Based on previous research, fetal overgrowth in the present study should be classified into pathological and non-pathological categories. Studies related to Large Offspring Syndrome (LOS) can be categorized as pathological, characterized not only by significantly increased birth weight but also by abnormal phenotypes such as inability to stand, respiratory distress, omphalocele, organomegaly, and asymmetric overgrowth [21]. Non-pathological overgrowth is mainly characterized by excessive birth weight, but the calves grow and develop normally after birth with overall good health, without the aforementioned malformations or dysfunctions. Although there are significant differences between the two in terms of calf health and disease risk, their impact on recipient cows in terms of dystocia and associated animal welfare issues is no different. Therefore, this study focused on calves with non-pathological overgrowth, aiming to explore its underlying mechanisms and influencing factors.

Compared with normal-weight (HCG) calves, the BIG group exhibited significantly elevated levels of GH, IGF-1, and NE, indicating that a systemic metabolic activation state provided an endocrine driving force for rapid fetal growth [22]. Meanwhile, the levels of ALB, ALT, TC, HDL, and LDL were also concurrently elevated, reflecting enhanced hepatic synthetic function and improved lipid transport capacity, which collectively established a metabolic basis supporting overgrowth [23, 24]. Notably, the levels of CRE, NEFA, and β-BHBA in the BIG group were significantly lower than those in the HCG group. Combined with the significant activation of the bile acid metabolism pathway, this indicates that the body primarily relied on exogenous lipid digestion and absorption rather than endogenous fat mobilization to meet energy demands. The reductions in creatinine and

ketone bodies further confirmed that renal load and hepatic fat mobilization were in a state of efficient compensation, effectively preventing the occurrence of metabolic disorders. This “exogenous supply-dominant” energy acquisition strategy suggests that non-pathological overgrowth did not exceed the body's tolerance threshold.

Metabolomics analysis further revealed significantly differentiated blood metabolic profiles between the BIG and HCG groups. KEGG enrichment pathways were predominantly concentrated on growth and development regulatory modules, including amino acid metabolism, insulin signaling, and nutrient digestion and absorption, rather than inflammatory pathways. Global activation of amino acid metabolism was a key feature underlying the formation of overgrowth, among which the arginine and proline metabolism pathway showed significant differential expression. Arginine is catalyzed by ARG1 to generate ornithine, which then participates in polyamine synthesis and creatine metabolism [25]; polyamines can activate mTOR signaling to stimulate protein synthesis, while the nitric oxide pathway enhances placental blood flow and nutrient transport efficiency [26]. The tryptophan metabolism pathway was also significantly enriched. In bovine endometrial epithelial cells, tryptophan stimulates IGFBP1 expression through the kynurenine pathway via the TDO2-AHR pathway, and IGFBP1 is a classical marker of uterine receptivity and embryonic elongation in ruminants, indicating that tryptophan metabolism also plays a regulatory role in embryo implantation and early development[27, 28]. The above amino acid metabolic reprogramming provides sufficient substrate and signal support for rapid fetal growth, constituting a core driving factor of overgrowth.

In addition to metabolic reprogramming, the adaptation of the immune and antioxidant systems also represents an important factor in the non-pathological phenotype. The levels of pro-inflammatory cytokines IL-6 and IL-2 in the BIG group were significantly elevated, but no corresponding mitochondrial dysfunction was observed, indicating that the inflammatory signals were within a controlled physiological range. Meanwhile, the levels of IgA and IgG were significantly reduced, which may reflect a reallocation of immune resources, whereby the body prioritizes limited resources for growth-related processes. In terms of antioxidant capacity, the activities of SOD, CAT, and GSH-Px as well as T-AOC levels in the BIG group were lower than those in the HCG group, but ROS levels also showed a decreasing trend. Given that SOD, CAT, and GSH-Px are the core enzymes for scavenging superoxide and peroxides, their reduced activities combined with decreased ROS levels suggest that the body did not suffer from antioxidant depletion, but rather was in an adaptive balance of “low ROS-low antioxidant enzymes”; the enrichment of the glutathione metabolism pathway further supports this view[29].

Transcriptomic and whole-transcriptome analyses systematically dissected the molecular mechanisms at the RNA regulatory level. mRNA differential expression and WGCNA analysis revealed that the ECM-receptor interaction, focal adhesion, PI3K-Akt, and cytokine-cytokine receptor pathways were significantly enriched, and the core hub genes (*ALB*, *IL1B*, *IL10*, *FGF2*, *HIF1A*) pointed to synergistic alterations in inflammatory immune regulation, angiogenesis, and oxygen homeostasis maintenance. Among these, activation of the ECM-receptor interaction pathway reflected enhanced extracellular matrix remodeling, while activation of the PI3K-Akt pathway

directly drove cell proliferation and metabolic regulation. lncRNA analysis screened 380 core differentially expressed lncRNAs, whose adjacent genes were enriched in metabolic pathways and immune-related pathways, suggesting that lncRNAs may participate in transcriptional reprogramming through cis-regulation [30]. In the ceRNA network constructed by miRNA and circRNA sequencing, two core axes, circ36279-bta-miR-431-*SPON2*/*TNN*/*INSIG1* and circ35151-bta-miR-154b-*SHD*/*TBX2*/*IL20RA*, were particularly prominent. INSIG1, as a negative regulator of cholesterol synthesis, may relieve the inhibition of cholesterol synthesis upon downregulation, thereby providing membrane lipid raw materials for rapid growth [31]; *SPON2*, as an extracellular matrix protein, plays an important role in placental vascular development [32]. The above ceRNA regulatory relationships indicate that multi-level synergistic regulation, ranging from lipid synthesis to angiogenesis, constitutes a post-transcriptional driving network for tissue overgrowth. Integrating metabolomic and transcriptomic data, the activation of metabolic and immune pathways at the mRNA level, cis-regulation by lncRNAs, and the deep regulation of lipid synthesis and angiogenesis signals by the ceRNA network collectively contribute to this regulatory landscape.

Given that both the BIG and HCG groups in this study consisted of non-pathological healthy calves, and considering practical production and animal welfare concerns, the measurements in this study mainly relied on peripheral blood samples. Blood can reflect systemic metabolic and immune status; however, this study still has limitations. For example, we failed to directly assess transcriptomic changes in key target organs such as the placenta, muscle, or liver, and the analyses were mainly based on high-throughput sequencing and bioinformatics correlation analysis. Future studies still need to conduct functional correlation analyses through validation in tissues and organs as well as cytological experiments.

In summary, through integrated multi-omics analysis, this study revealed the molecular mechanisms underlying the multi-level synergistic actions in non-pathological overgrowth calves derived from *in vitro* embryos, encompassing hormone driving, amino acid metabolic reprogramming, adaptive remodeling of immune and antioxidant systems, and post-transcriptional regulation. Based on these findings, in future IVP practice, efforts can be made to optimize the concentration ratios of key amino acids such as arginine and tryptophan in culture media, or to target relevant metabolic pathways (e.g., polyamine synthesis, NO production, and bile acid metabolism), aiming to reduce the risk of abnormal overgrowth while ensuring embryo development efficiency, thereby providing a theoretical basis and potential intervention targets for improving the IVP technical system.

Conclusions

This study investigated the molecular regulatory network of non-pathological fetal overgrowth in IVP Holstein calves through integrated multi-omics analysis. Physiological, biochemical, and hormonal parameters showed significant changes, but no pathological alterations were observed. Two ceRNA regulatory axes, circ36279-bta-miR-431 (targeting *SPON2*, *TNN*, *INSIG1*, and *AEN*) and circ35151-bta-miR-154b (targeting *SHD*, *TBX2*, and *IL20RA*), were identified as key regulatory

nodes. Integrative multi-omics analysis further pinpointed the arginine and proline metabolism pathway as a core metabolic hub, in which *AGMAT*, *ARG1*, and *CKM* serve as rate-limiting enzymes driving the downstream flux of multiple metabolic branches, including polyamine, ornithine, creatine, and proline metabolism. In summary, this study reveals a molecular network in which non-coding RNAs and metabolites synergistically regulate the fetal overgrowth phenotype, providing data-based evidence and theoretical references for optimizing the IVP production system and developing precise intervention strategies for abnormal birth weight in calves.

Abbreviations

ALB	Albumin
ALT	Alanine aminotransferase
β-BHBA	β-Hydroxybutyrate
BDNF	Brain-derived neurotrophic factor
BIG	Big group
BP	Biological process
BSJ	Back-splicing junction
CAT	Catalase
CC	Cellular component
ceRNA	Competing endogenous RNA
circRNA	Circular RNA
CRE	Creatinine
DEC	Differentially expressed circRNA
DEG	Differentially expressed gene
DEL	Differentially expressed lncRNA
DEM	Differentially expressed miRNA
DG	Diacylglycerol
ECM	Extracellular matrix

ET	Embryo transfer
FDR	False discovery rate
GH	Growth hormone
GO	Gene Ontology
GS	Gene significance
GSH-Px	Glutathione peroxidase
HCG	Healthy control group
HDL	High-density lipoprotein cholesterol
IGF-1	Insulin-like growth factor 1
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IL	Interleukin
IQR	Interquartile range
IVF	*In vitro* fertilization
IVM	*In vitro* maturation
IVP	*In vitro* embryo production
KEGG	Kyoto Encyclopedia of Genes and Genomes
LA	Linoleic acid
LDL	Low-density lipoprotein cholesterol
lncRNA	Long non-coding RNA
LOS	Large offspring syndrome
LPC	Lysophosphatidylcholine
MAD	Median absolute deviation
ME	Module eigengene
MF	Molecular function

miRNA	MicroRNA
MM	Module membership
mRNA	Messenger RNA
NE	Norepinephrine
NEFA	Non-esterified fatty acids
OPU	ovum pick-up
PC	Phosphatidylcholine
PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
PPI	Protein-protein interaction
RPM	Reads per million mapped reads
ROS	Reactive oxygen species
SD	Standard deviation
SOD	Superoxide dismutase
T-AOC	Total antioxidant capacity
TC	Total cholesterol
TG	Triglyceride
TPM	Transcripts per million
TMR	Total mixed ration
WGCNA	Weighted gene co-expression network analysis

Supplementary Information

Acknowledgements

We gratefully acknowledge Professor Huabin Zhu from the Institute of Animal Science, Chinese Academy of Agricultural Sciences, for their assistance with data analysis and professional support.

Authors' contributions

Leiyu Qu: Investigation, Data curation, Formal analysis, Visualization, Writing - original draft. Zhengyu Hang: Writing - original draft, Writing - review & editing. Lei Wang: Formal analysis, Writing - review & editing. Yao Chang: Conceptualization, Methodology. Ruina Ma: editing. ShanJiang Zhao: Investigation. Qingshan Gao: review & editing. Qing Xu: Writing - review & editing. Yachun Wang: Conceptualization, Methodology, Project administration, Supervision.

Funding

This study was funded by the China National Modern Agricultural Industry Technology System (project number: CARS-35).

Declarations

Ethics approval and consent to participate

All experimental procedures complied with the guidelines of the Animal Ethics Committee of China Agricultural University (Beijing, China) (Approval code: AW70505202-1-01).

Clinical trial number:

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the Genome Sequence Archive (GSA) repository, [PRJCA070992].

Competing interests

The authors declare that they have no competing interests.

References

1. Figueiredo JHMVBDMSRM de MLFRFRA. Oocyte developmental potential and embryo production before puberty in cattle. Oocyte developmental potential and embryo production before puberty in cattle. 2024;21:0–0. https://doi.org/10.1590/1984-3143-AR2024-0069.

2. Viana JH. 2023 Statistics of embryo production and transfer in domestic farm animals.

3. Rivera RM, Goldkamp AK, Patel BN, Hagen DE, Soto-Moreno EJ, Li Y, et al. Identification of large offspring syndrome during pregnancy through ultrasonography and maternal blood transcriptome analyses. Sci Rep. 2022;12:10540. https://doi.org/10.1038/s41598-022-14597-w.

4. Murphy EM, Doyle RC, Lonergan P, Butler ST. Gestation length, calving characteristics, and early life health of calves born following transfer of fresh in vitro-produced embryos generated using conventional or sex-sorted sperm. Journal of Dairy Science. 2026;:S0022030226028730. https://doi.org/10.3168/jds.2026-28566.

5. Rabaglino MB, Bojsen-Møller Secher J, Sirard M, Hyttel P, Kadarmideen HN. Epigenomic and transcriptomic analyses reveal early activation of the HPG axis in in vitro-produced male dairy calves. The FASEB Journal. 2021;35:e21882. https://doi.org/10.1096/fj.202101067R.

6. Zhang X, Li F, Peng L, Huang W, Du Y, Yang L, et al. Integrated multi-omics analysis of metabolome and transcriptome profiles during bovine adipocyte differentiation reveals functional divergence of FADS2 isoforms in lipid metabolism regulation. BMC Genomics. 2025;26:457. https://doi.org/10.1186/s12864-025-11650-6.

7. Xu H, Gong L, Guo Y. Integrated transcriptomics and metabolomics reveal multi-target mechanisms of tannins against Clostridium perfringens and necrotic enteritis. J Animal Sci Biotechnol. 2025;16:98. https://doi.org/10.1186/s40104-025-01228-3.

8. Xu C, Shah MA, Mipam T, Wu S, Yi C, Luo H, et al. Bovid microRNAs involved in the process of spermatogonia differentiation into spermatocytes. Int J Biol Sci. 2020;16:239–50. https://doi.org/10.7150/ijbs.38232.

9. Zhou L, Chen J, Li Z, Li X, Hu X, Huang Y, et al. Integrated Profiling of MicroRNAs and mRNAs: MicroRNAs Located on Xq27.3 Associate with Clear Cell Renal Cell Carcinoma. PLoS ONE. 2010;5:e15224. https://doi.org/10.1371/journal.pone.0015224.

10. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11:1650–67. https://doi.org/10.1038/nprot.2016.095.

11. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495:333–8. https://doi.org/10.1038/nature11928.

12. Gao Y, Zhang J, Zhao F. Circular RNA identification based on multiple seed matching. Briefings in Bioinformatics. 2018;19:803–10. https://doi.org/10.1093/bib/bbx014.

13. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Research. 2019;47:D155–62. https://doi.org/10.1093/nar/gky1141.

14. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Research. 2012;40:37–52. https://doi.org/10.1093/nar/gkr688.

15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8.

16. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Research. 2011;39 suppl_2:W316–22. https://doi.org/10.1093/nar/gkr483.

17. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003;5:R1. https://doi.org/10.1186/gb-2003-5-1-r1.

18. Kruger J, Rehmsmeier M. RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Research. 2006;34 Web Server:W451–4. https://doi.org/10.1093/nar/gkl243.

19. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.

20. Rivera RM, Goldkamp AK, Patel BN, Hagen DE, Soto-Moreno EJ, Li Y, et al. Identification of large offspring syndrome during pregnancy through ultrasonography and maternal blood transcriptome analyses. Sci Rep. 2022;12:10540. https://doi.org/10.1038/s41598-022-14597-w.

21. Nava-Trujillo H, Rivera RM. Review: Large offspring syndrome in ruminants: current status and prediction during pregnancy. animal. 2023;17:100740. https://doi.org/10.1016/j.animal.2023.100740.

22. Butler ST, Sitko EM, Lucy MC. The impact of cow genetics on the somatotropic axis in pasture-based systems: implications for production and reproduction. JDS Communications. 2026;:S2666910226000608. https://doi.org/10.3168/jdsc.2025-0985.

23. Sato H, Kumano R, Fukumori R, Oikawa S. Lipoprotein composition of calves before and after weaning and comparison with adult cows. J Vet Med Sci. 2025;87:167–70. https://doi.org/10.1292/jvms.24-0172.

24. Li C, Huang J, Chen X, Yan Y, Li L, Zhao W. Transcriptome Analysis Reveals That NEFA and β-Hydroxybutyrate Induce Oxidative Stress and Inflammatory Response in Bovine Mammary Epithelial Cells. Metabolites. 2022;12:1060. https://doi.org/10.3390/metabo12111060.

25. Wu G, Bazer FW, Satterfield MC, Li X, Wang X, Johnson GA, et al. Impacts of arginine nutrition on embryonic and fetal development in mammals. Amino Acids. 2013;45:241–56. https://doi.org/10.1007/s00726-013-1515-z.

26. Halloran KM, Stenhouse C, Wu G, Bazer FW. Arginine, Agmatine, and Polyamines: Key Regulators of Conceptus Development in Mammals. In: Wu G, editor. Amino Acids in Nutrition and Health. Cham: Springer International Publishing; 2021. p. 85–105. https://doi.org/10.1007/978-3-030-74180-8_6.

27. Simmons RM, Erikson DW, Kim J, Burghardt RC, Bazer FW, Johnson GA, et al. Insulin-Like Growth Factor Binding Protein-1 in the Ruminant Uterus: Potential Endometrial Marker and Regulator of Conceptus Elongation. Endocrinology. 2009;150:4295–305. https://doi.org/10.1210/en.2009-0060.

28. Wang P-C, Liu Z-K, Li J-R, Zhao Z-H, Chang Q-W, Guo X-M, et al. Tryptophan regulates the expression of IGFBP1 in bovine endometrial epithelial cells in vitro via the TDO2-AHR pathway. BMC Vet Res. 2024;20:390. https://doi.org/10.1186/s12917-024-04191-9.

29. Nejad FM, Mohammadabadi M, Roudbari Z, Gorji AE, Sadkowski T. Network visualization of genes involved in skeletal muscle myogenesis in livestock animals. BMC Genomics. 2024;25:294. https://doi.org/10.1186/s12864-024-10196-3.

30. Lu J, Wu J, Zhao Z, Wang J, Chen Z. Circulating LncRNA Serve as Fingerprint for Gestational Diabetes Mellitus Associated with Risk of Macrosomia. Cell Physiol Biochem. 2018;48:1012–8. https://doi.org/10.1159/000491969.

31. Suzuki A, Ogata K, Yoshioka H, Shim J, Wassif CA, Porter FD, et al. Disruption of Dhcr7 and Insig1/2 in cholesterol metabolism causes defects in bone formation and homeostasis through primary cilium formation. Bone Res. 2020;8:1. https://doi.org/10.1038/s41413-019-0078-3.

32. Ghafouri F, Dehghanian Reyhan V, Sadeghi M, Miraei-Ashtiani SR, Kastelic JP, Barkema HW, et al. Competing Endogenous RNAs (ceRNAs) and Application of Their Regulatory Networks in Complex Traits and Diseases of Ruminants. Ruminants. 2024;4:165–81. https://doi.org/10.3390/ruminants4020011.

Figures

Figure 1. Comparisons of phenotypic traits between the HCG control group and BIG treatment group

(A)Birth weight of offspring calves in the HCG control group and BIG treatment group. Significant difference was detected between the two groups (**, $P < 0.01$).(B) Gestation days of recipient cows between the HCG and BIG groups. A significant difference was observed (*, $P < 0.05$).(C) Postpartum weight of recipient cows in the two groups; no significant difference (ns, $P \geq 0.05$) was

found.(D) Postpartum height of recipient cows between HCG and BIG groups, with a significant difference (*, $P < 0.05$) identified.

Figure 2. Differences in blood physiological, biochemical, and hormonal indicators between HCG control group and BIG treatment group.

The horizontal line inside each box represents the median value, the box spans the interquartile range (IQR, 25th–75th percentiles), and the whiskers extend to the minimum and maximum values within 1.5×IQR. The violin shapes illustrate the distribution density of the data, with individual data points overlaid. *$P < 0.05$; **$P <0.01$.

Figure 3. Metabolomic analysis of HCG and BIG calves.

(A)Pie chart showing the chemical classification of detected metabolites. (B) PCA score plot showing the overall separation of metabolite profiles between the two groups. (C) Venn diagram of metabolites identified in the two groups. (D) Volcano plot of differential metabolites. (E) KEGG pathway enrichment analysis of differential metabolites.

Figure 4. Differential mRNA transcriptome analysis between HCG control group and BIG treatment group.

(A)Principal component analysis (PCA) plot showing mRNA expression profiles. Red dots represent HCG group, while green dots represent calves with BIG group.(B) Volcano plot of differentially expressed genes. Blue dots indicate downregulated genes, and red dots indicate upregulated genes in BIG calves. (C) Protein-protein interaction (PPI) network analysis of differentially expressed genes. Node connectivity was calculated using the Degree algorithm in the CytoHubba plugin, and genes with the highest Degree values were identified as candidate hub genes. Node color and size were mapped to Degree values, with darker and larger nodes representing higher connectivity. (D) Soft-thresholding power analysis for WGCNA of mRNA expression profiles. The scale-free fitting index (R2) is plotted against soft-thresholding power (1-30). Horizontal dashed lines mark R2 thresholds of 0.8, 0.7, and 0.6. (E)Module-trait relationships in WGCNA of mRNA. Each row is a module eigengene; each column is a group (HCG or BIG). The heatmap shows the Pearson correlation coefficient (top) and *P*-value (in parentheses). Red indicates positive correlation, blue indicates negative correlation.(F) GO functional and KEGG pathway enrichment analyses of differentially expressed genes

Figure 5. WGCNA and association analysis based on lncRNA expression.

(A)PCA plot of lncRNA expression profiles.(B) Volcano plot of differentially expressed lncRNAs.(C) Soft-thresholding power selection for WGCNA.(D) Heatmap of module-trait relationships between WGCNA modules and groups.(E) Venn diagram showing the overlap between differentially expressed lncRNAs and those from significant WGCNA modules.(F) KEGG pathway enrichment of genes within 100 kb of overlapping lncRNAs.

Figure 6. Multi-omics expression profiling and ceRNA network analysis between HCG and BIG groups

(A)PCA plot of miRNA expression profiles, with red dots for HCG group and green dots for BIG group.(B) Volcano plot of differentially expressed miRNAs.(C) PCA plot of circRNA expression profiles.(D) Volcano plot of differentially expressed circRNAs, coloring criteria consistent with panel B.(E) Schematic diagrams of two predicted circRNA-miRNA-mRNA ceRNA regulatory networks.(F) The Venn diagram illustrates the overlapping KEGG pathways identified by metabolomic and transcriptomic analyses.

Figure 7 Heatmap showing correlations between candidate genes and metabolites

Pearson correlation coefficients between candidate genes and metabolites are visualized in the heatmap. Red represents positive correlations and blue represents negative correlations. Values inside each cell correspond to correlation coefficients, with darker colors reflecting greater correlation strength. Asterisks (*, **) denote different statistical significance levels.

Figure 8. Arginine and proline metabolism pathway-related expression changes.

The figure displays a schematic of the Arginine and proline metabolism pathway (left) and the corresponding gene expression heatmap, illustrating differences in pathway-related gene expression between healthy (HCG) and BIG calves.

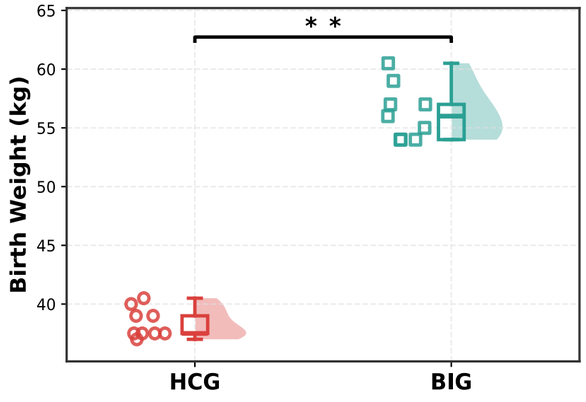
Tables

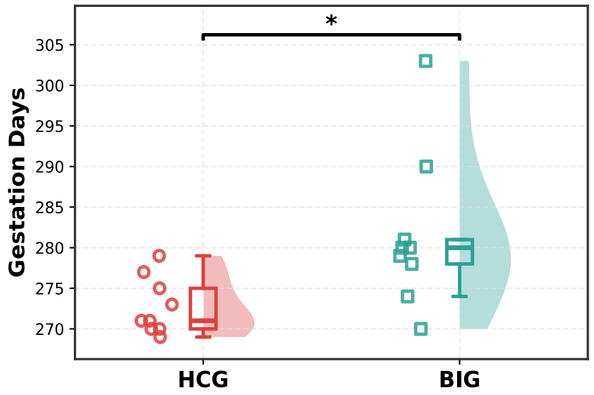
Table 1. Differentially expressed genes and differential metabolites in overlapping KEGG pathways between transcriptomic and metabolomic analyses.

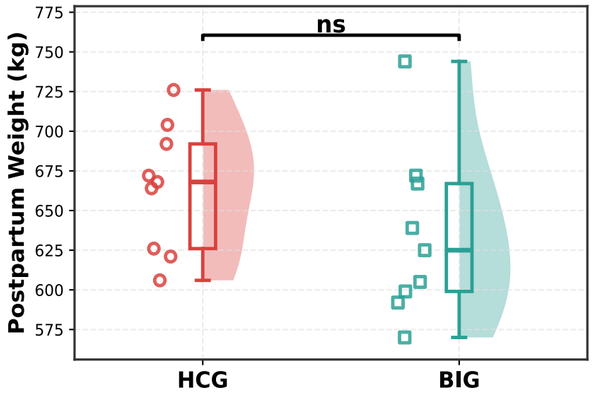
Pathway Name	Differentially Expressed Genes	Differential Metabolites
Arginine and proline metabolism	CKMT2,CKM,MAOB,ARG1,P4HA3,HOGA1,AGMAT	L-Glutamic Gamma-Semialdehyde|D-Octopine|Creatinine|L-Proline|Pyruvic Acid|Trans-4-Hydroxy-L-Proline|4-Acetamidobutyric Acid|Gamma-Aminobutyric Acid|1-Pyrroline-5-Carboxylic Acid|(3R,5S)-1-Pyrroline-3-Hydroxy-5-Carboxylic Acid

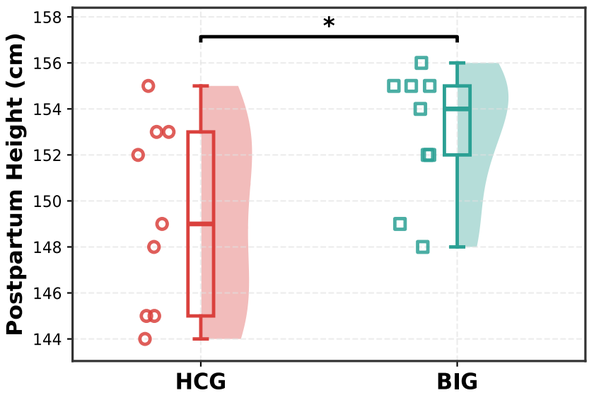
Tryptophan metabolism	TPH2,DDC,TPH1,AANAT,MAOB, AOX1,IDO1	5-Methoxyindoleacetic Acid|Serotonin|3-Hydroxy-3-Methyl-2,3-Dihydro-1H-Indol-2-One|3-Indoleacetic Acid|5-Hydroxy-N-Formylkynurenine|3-Indoleacetonitrile|5-Hydroxyindole-3-Acetic Acid|Formyl-5-Hydroxykynurenamine|Indole-3-Glycol Aldehyde
Neuroactive ligand-receptor interaction	GRIA1,NPB,NPFFR1,MLNR,OXT, ADRB2,MCHR1,LYPD6,GHR,GR M2,GABRR1,C5,EDNRB,P2RY4,C HRNG,TAC3,HCRT,INSL3,DRD3,P RSS2,GRIA3,DRD5,GABRQ,GHSR ,GZMA,SSTR1,F2,GRIN2C,PRLR,S LURP1,CORT,LTB4R2,MC5R,GHR L	L-Aspartic Acid|Serotonin|Adenosine 5'-Diphosphate|N-Acetyl-1-Aspartylglutamic Acid|L-Glycine|Gamma-Aminobutyric Acid
Protein digestion and absorption	SLC36A3,CPA1,COL23A1,COL11 A2,ATP1A4,ATP1A3,ACE2,MEP1B ,COL4A2,KCNQ1,FXYD2,COL5A3 ,COL4A3,COL6A1,COL9A3,COL6 A6,PRSS2	L-Aspartic Acid|L-Glycine|L-Proline|L-Alanine|L-Cysteine
Mineral absorption	SLC34A3,S100G,TRPV6,SLC34A1, FXYD2,ATP1A4,ATP1A3,ATP2B2	L-Glycine|L-Proline|L-Alanine

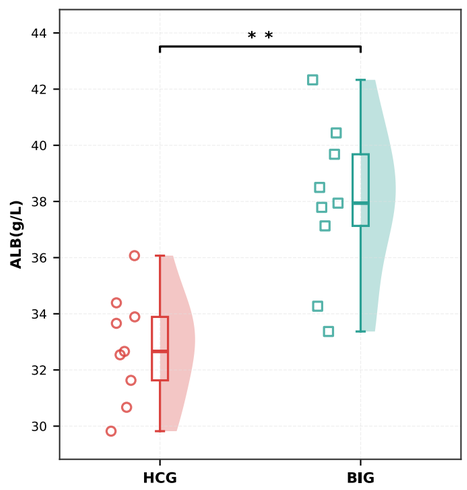
691

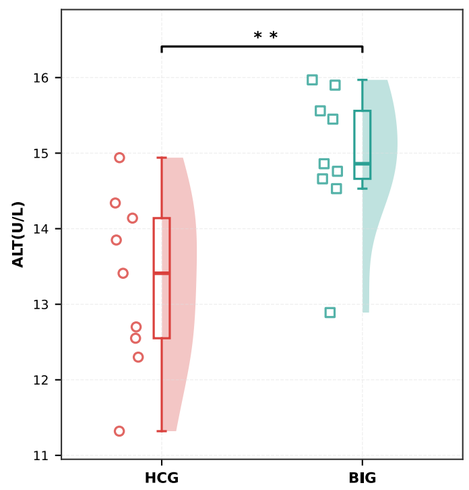

Birth Weight (kg)
65
60
55
50
45
40
* *
HCG
BIG

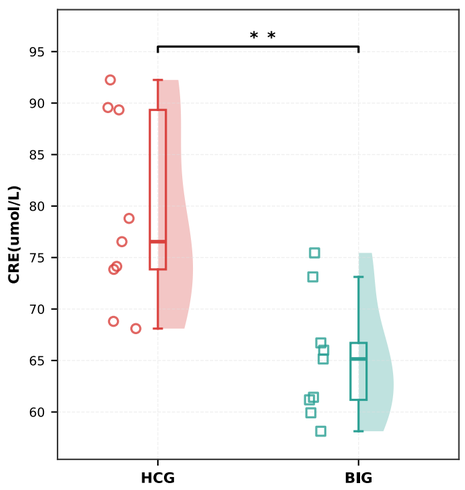

Gestation Days
305
300
295
290
285
280
275
270
*
HCG
BIG

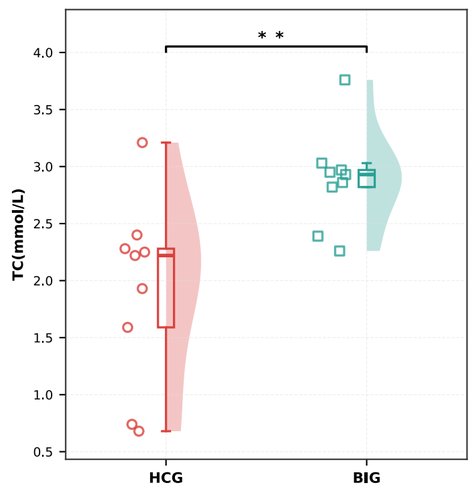

Postpartum Weight (kg)
775
750
725
700
675
650
625
600
575
ns
HCG
BIG

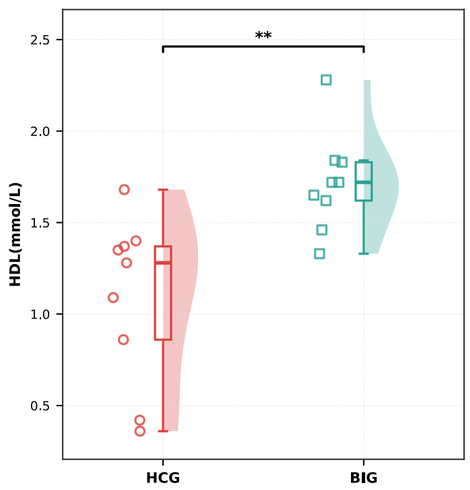

Postpartum Height (cm)
158
156
154
152
150
148
146
144
*
HCG
BIG

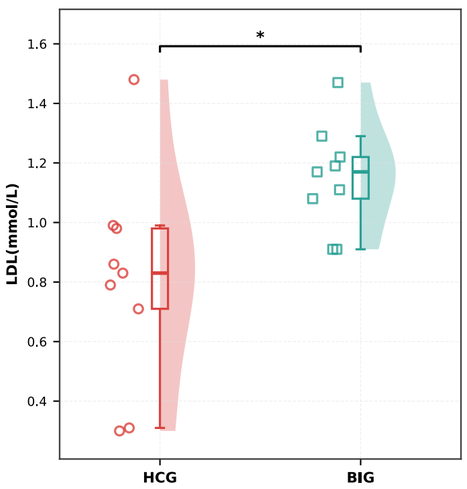

ALB(g/L)
HCG
BIG
* *

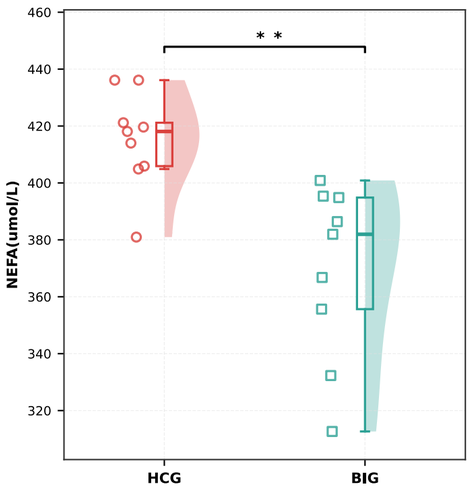

ALT(U/L)
HCG
BIG
* *

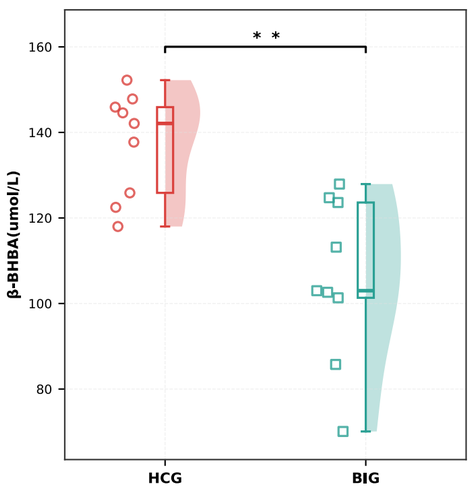

CRE(umol/L)
HCG
BIG
* *

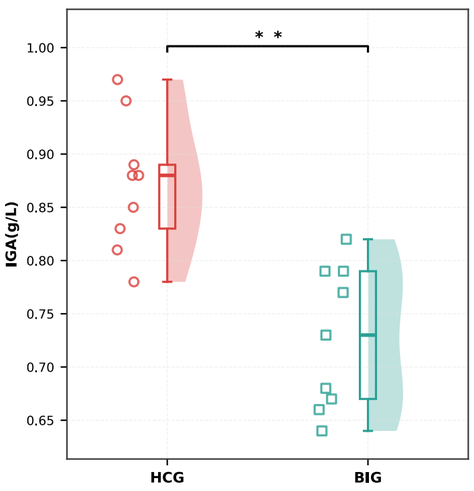

TC(mmol/L)
HCG
BIG
* *

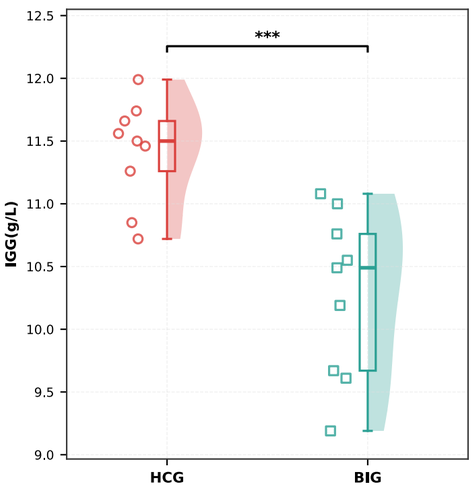

HDL(mmol/L)
HCG
BIG
**

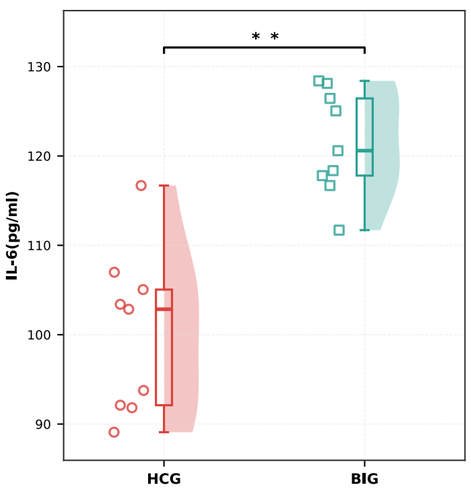

LDL(mmol/L)
HCG
BIG
*

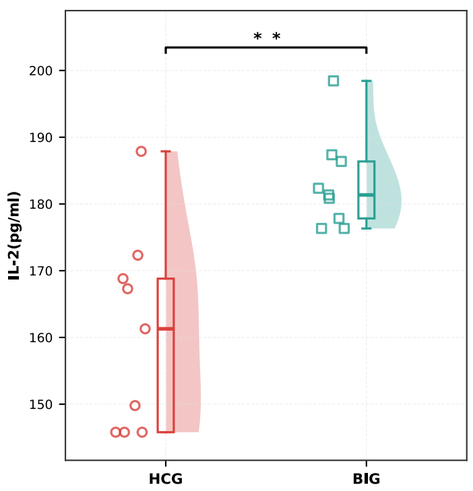

NEFA(umol/L)
HCG
BIG
* *

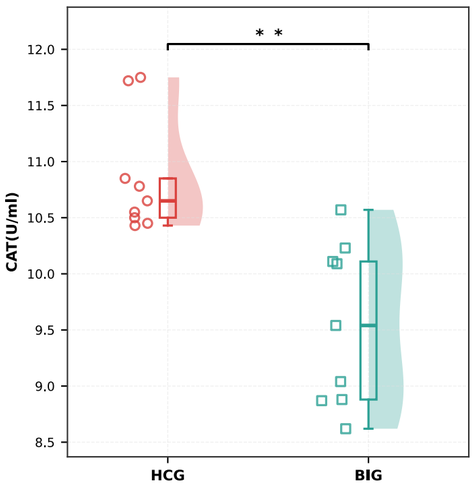

β-BHBA(umol/L)
HCG
BIG
* *

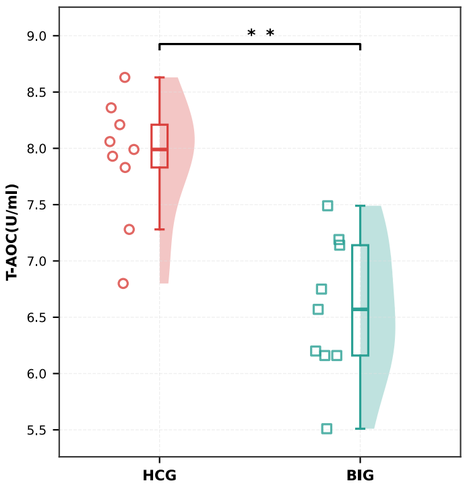

IGA(g/L)
HCG
BIG
* *

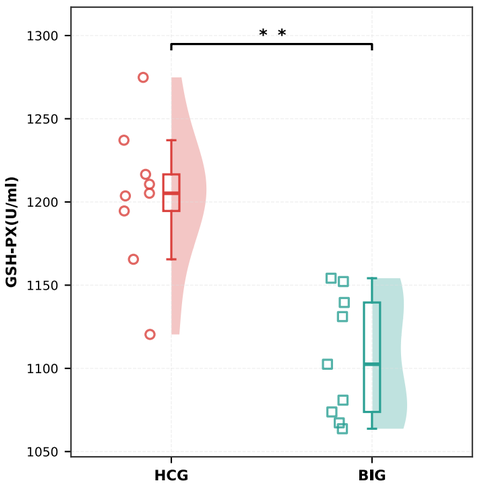

IGG(g/L)
HCG
BIG
***

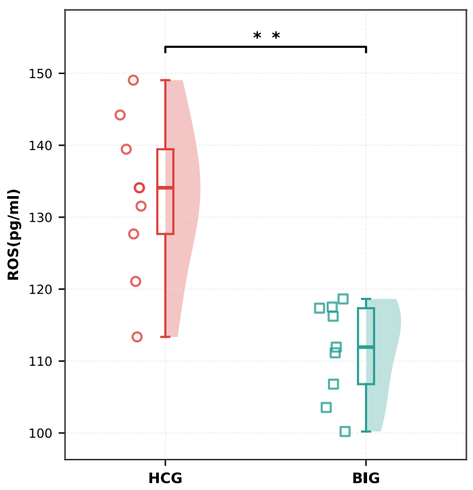

IL-6(pg/ml)
HCG
BIG
* *

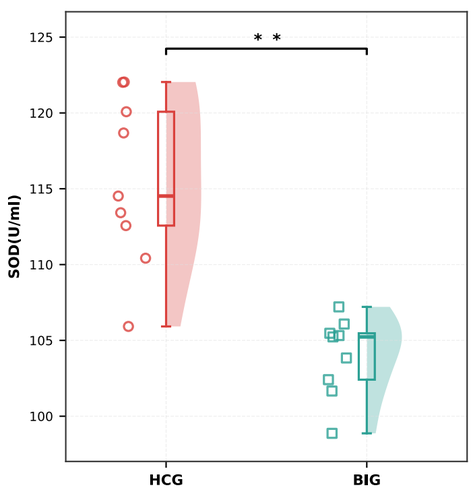

IL-2(pg/ml)
HCG
BIG
* *

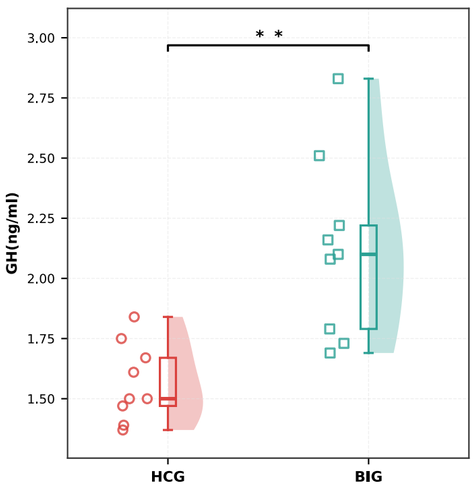

CAT(U/ml)
HCG
BIG
* *

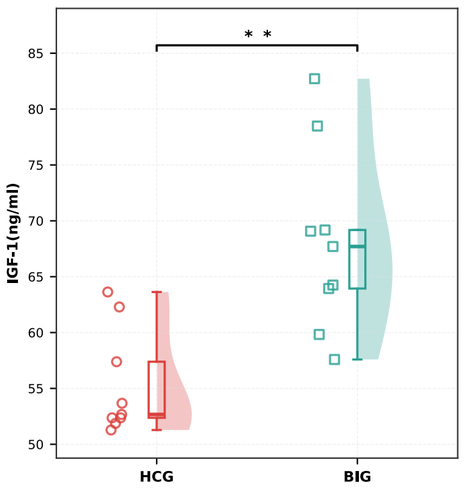

T-AOC(U/ml)
HCG
BIG
* *

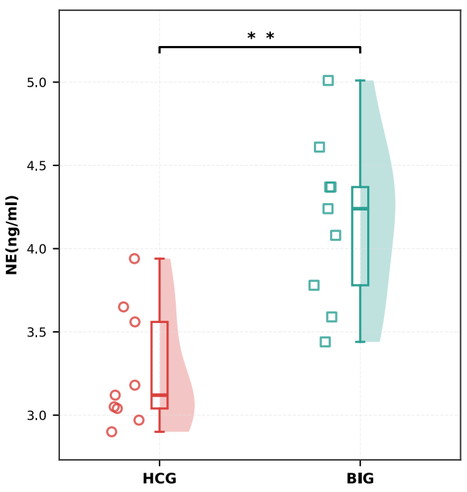

GSH-PX(U/ml)
HCG
BIG
* *


ROS(pg/ml)
HCG
BIG
* *

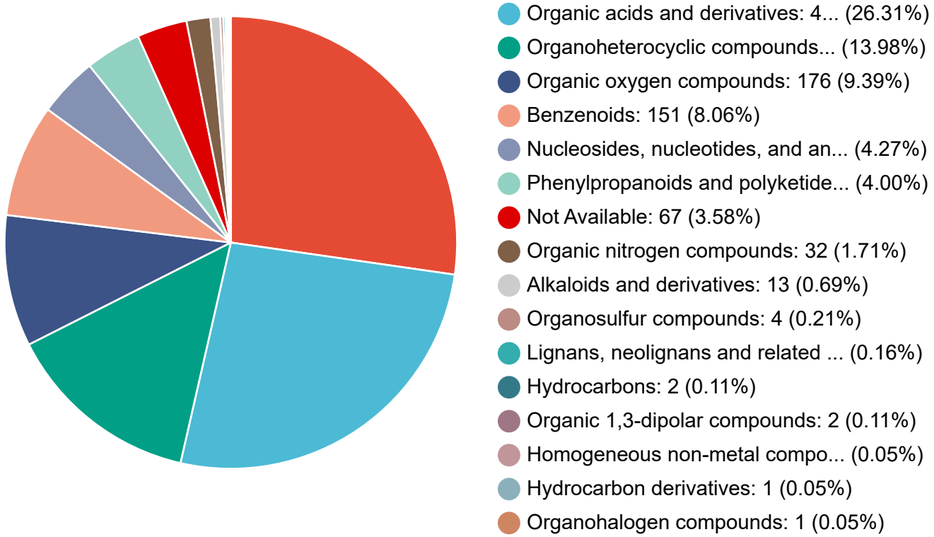

SOD(U/ml)
HCG
BIG
* *


GH(ng/ml)
HCG
BIG
* *

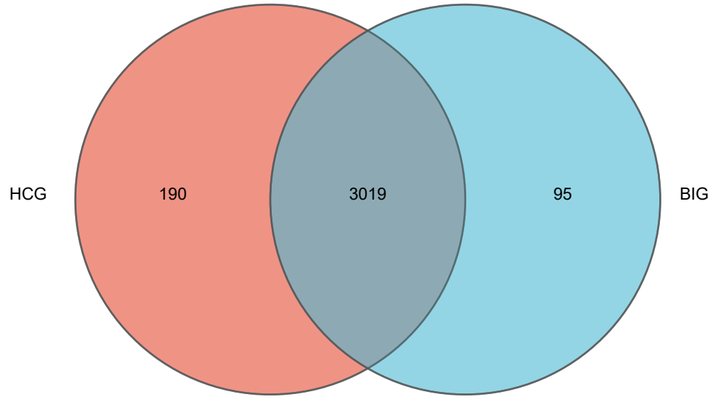

IGF-1(ng/ml)
HCG
BIG
* *


NE(ng/ml)
HCG
BIG
* *

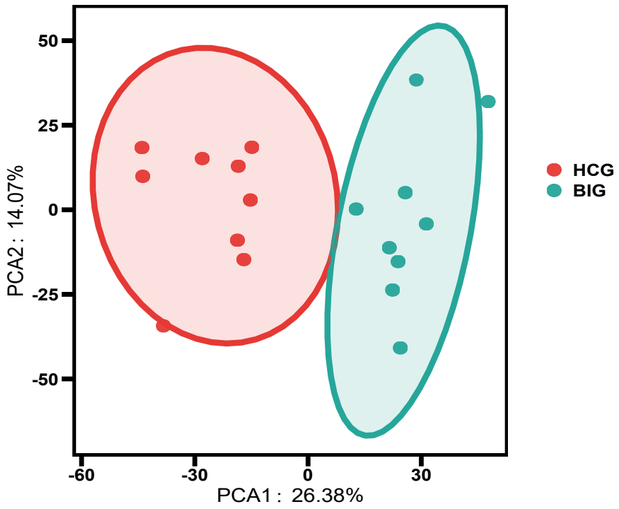
A.


Lipids and lipid-like molecules: ... (27.27%)
Organic acids and derivatives: 4... (26.31%)
Organoheterocyclic compounds... (13.98%)
Organic oxygen compounds: 176 (9.39%)
Benzenoids: 151 (8.06%)
Nucleosides, nucleotides, and an... (4.27%)
Phenylpropanoids and polyketide... (4.00%)
Not Available: 67 (3.58%)
Organic nitrogen compounds: 32 (1.71%)
Alkaloids and derivatives: 13 (0.69%)
Organosulfur compounds: 4 (0.21%)
Lignans, neolignans and related ... (0.16%)
Hydrocarbons: 2 (0.11%)
Organic 1,3-dipolar compounds: 2 (0.11%)
Homogeneous non-metal compo... (0.05%)
Hydrocarbon derivatives: 1 (0.05%)
Organohalogen compounds: 1 (0.05%)

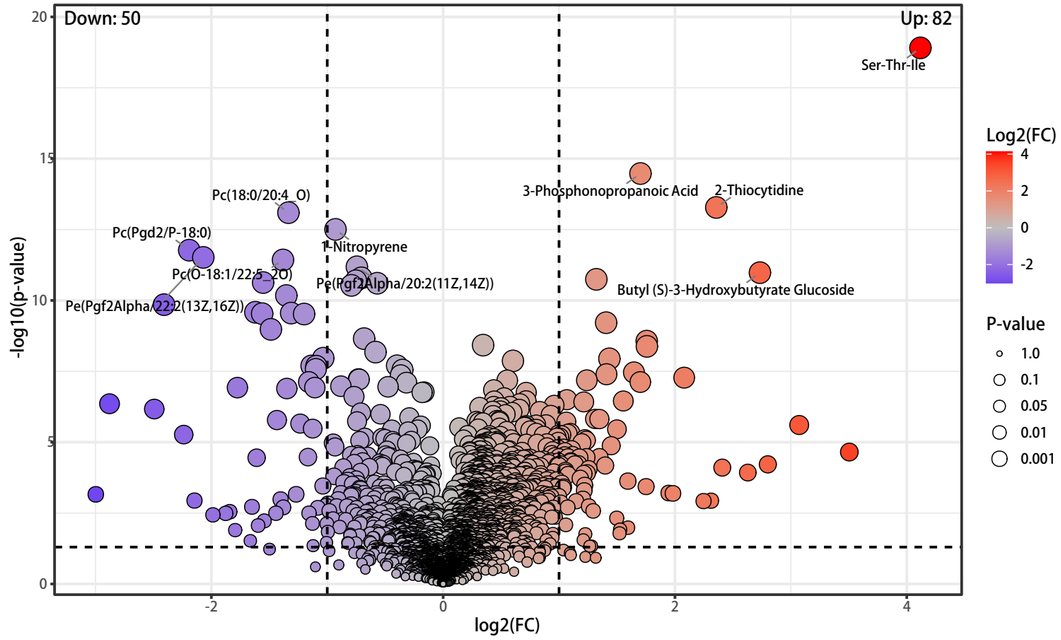
B.


HCG
190
3019
95
BIG

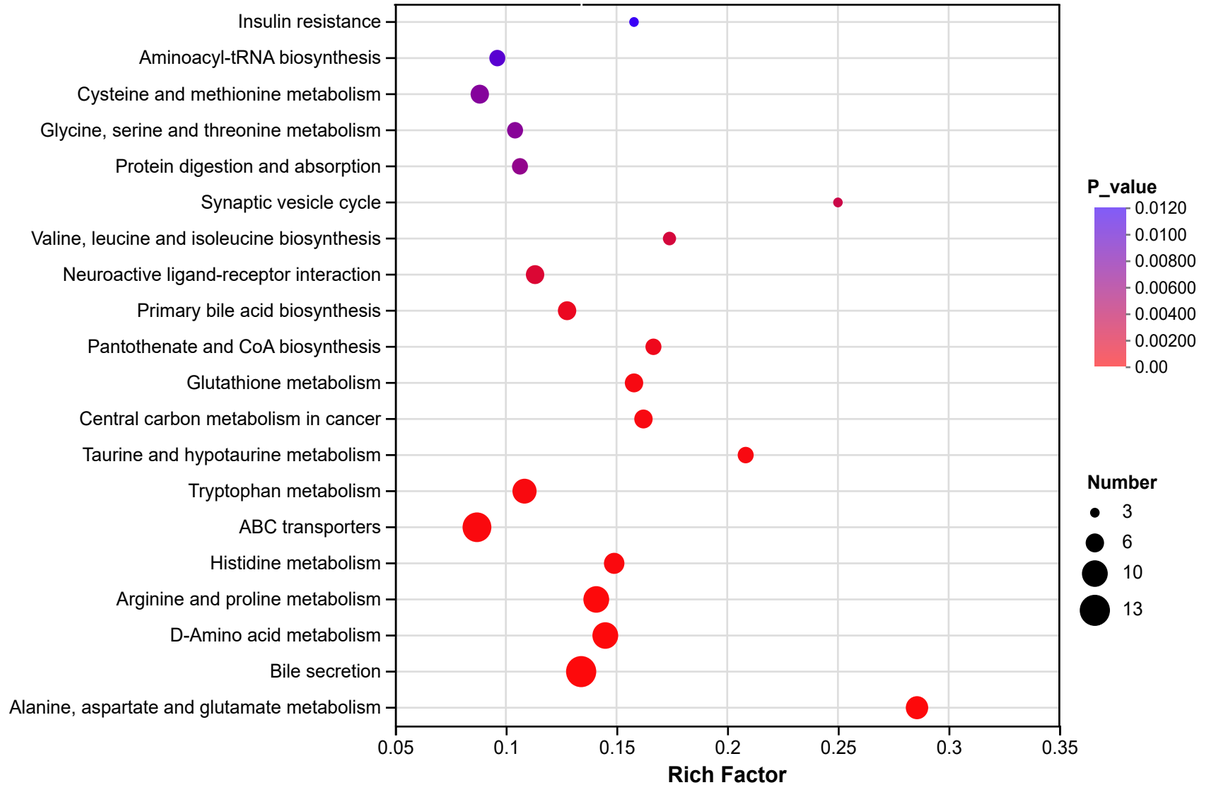
C.

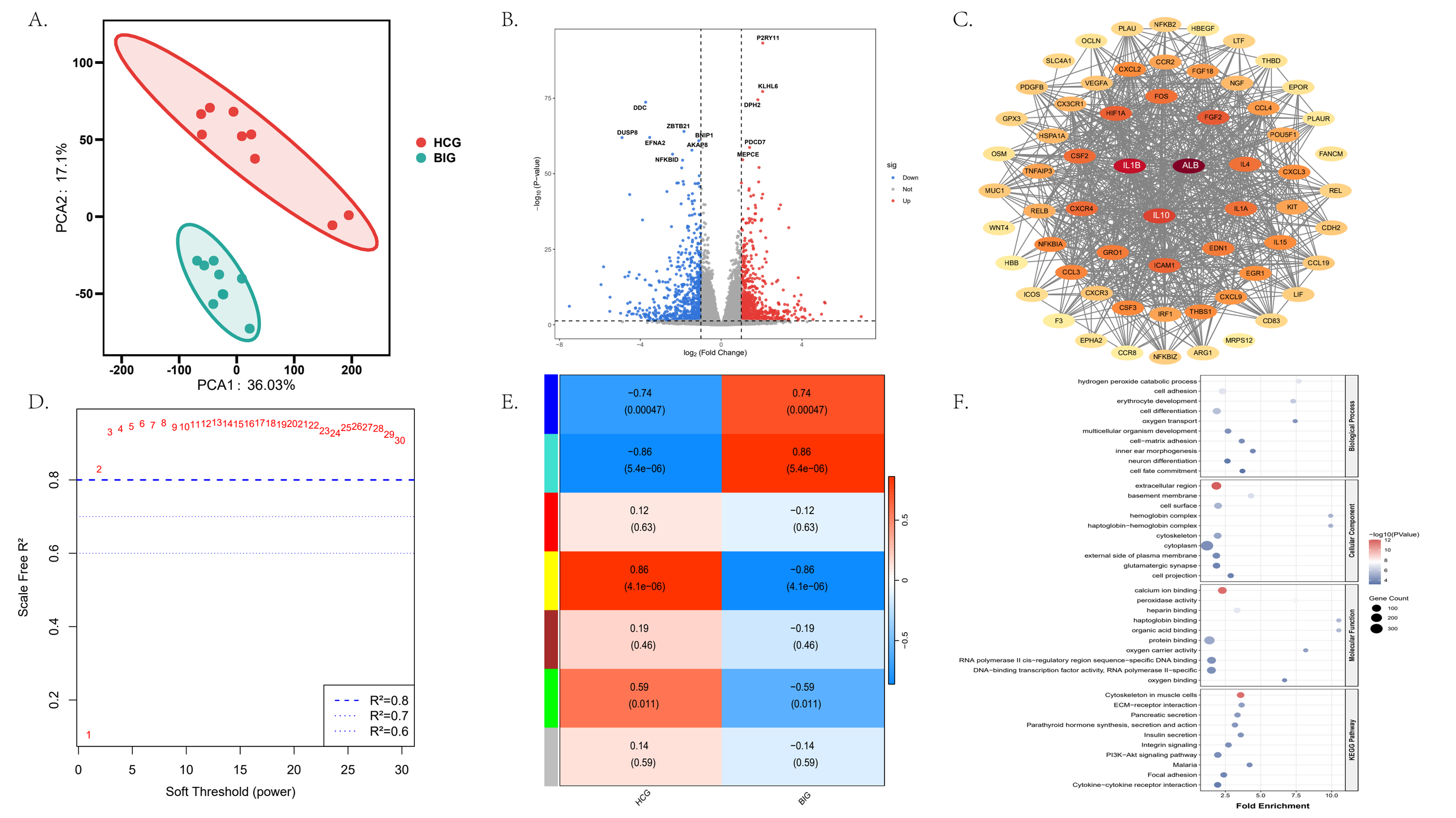

50
25
0
-25
-50
PCA2 : 14.07%
-60
-30
0
30
PCA1 : 26.38%
HCG
BIG

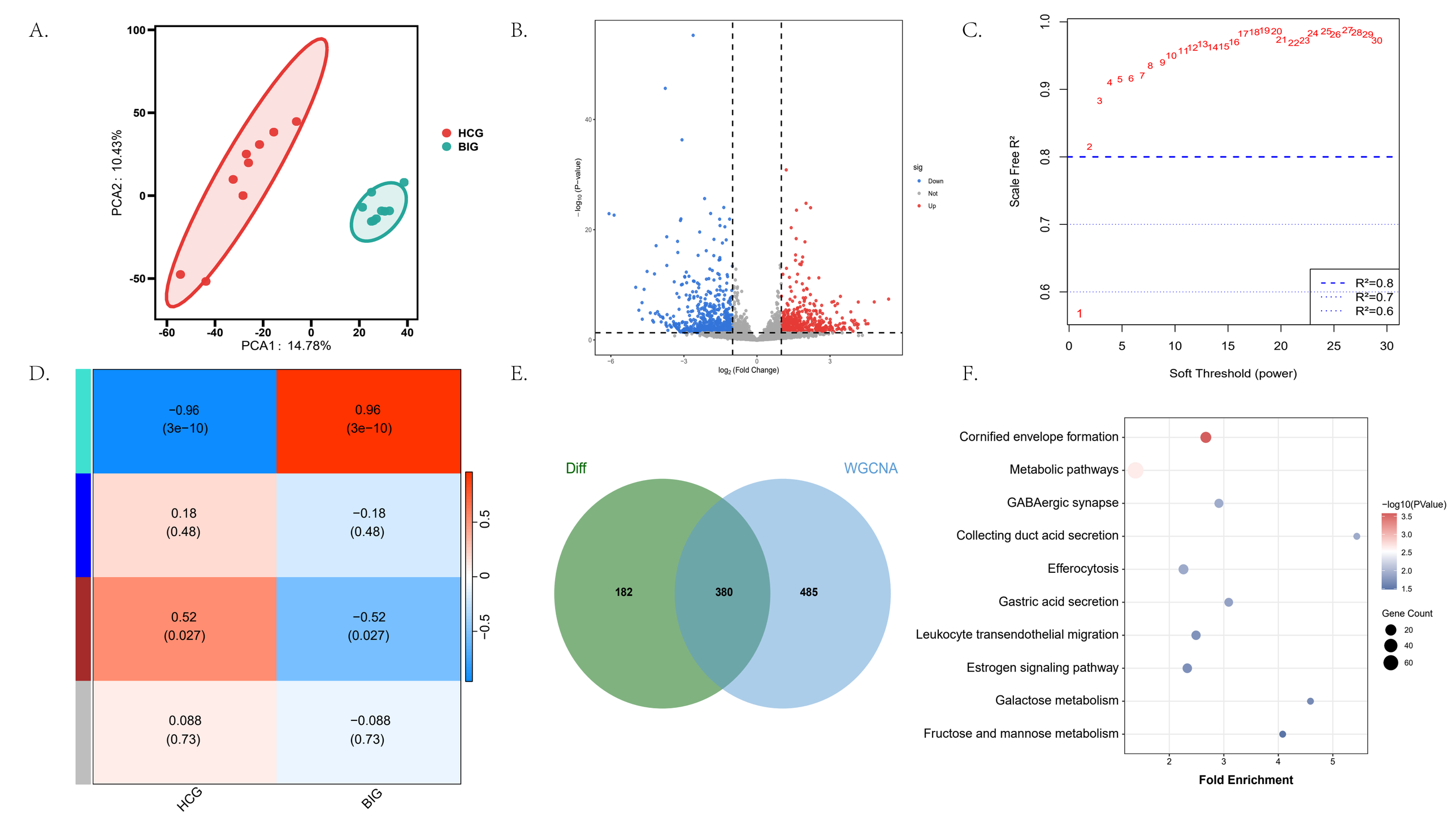
D.

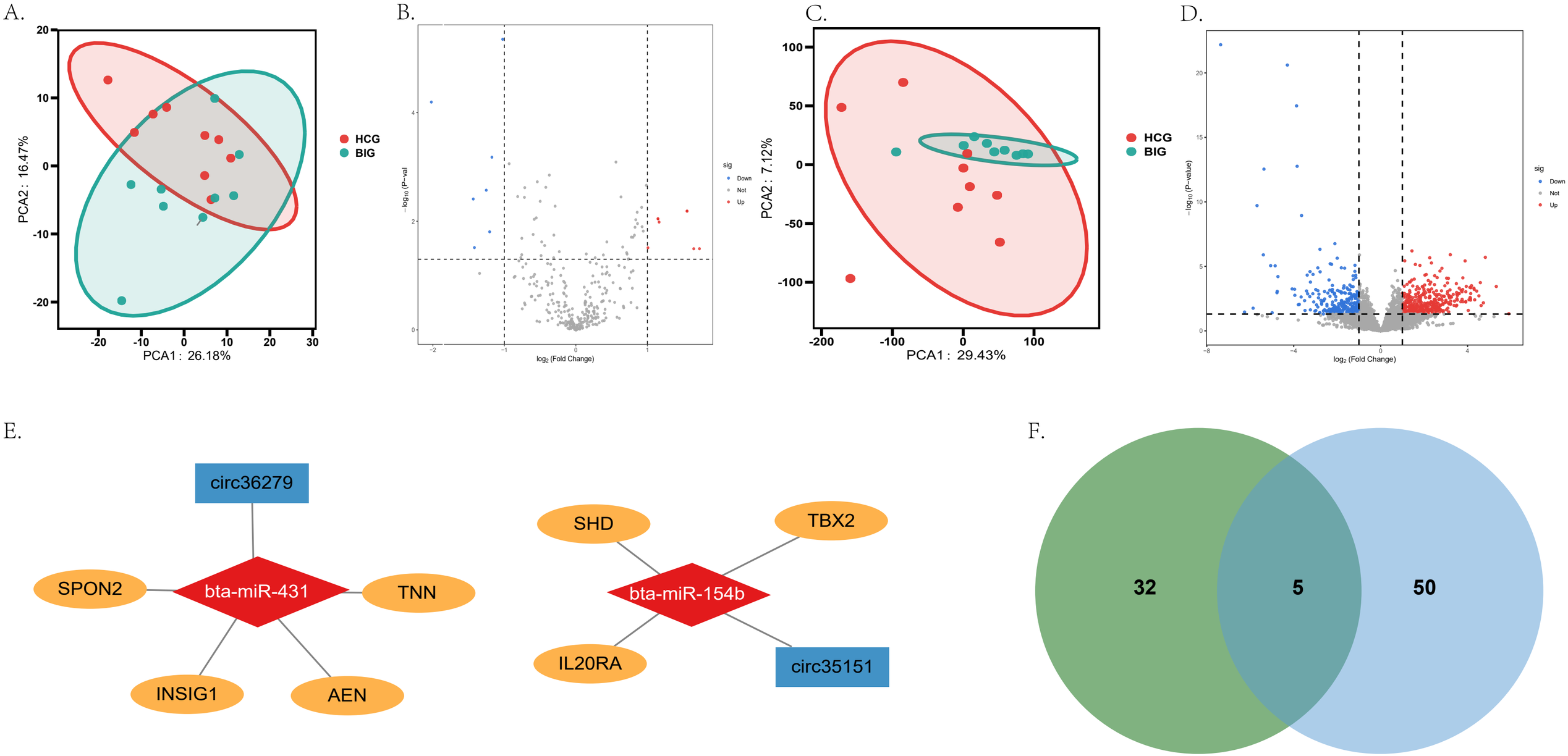

Down: 50
Up: 82
Ser-Thr-Ile
3-Phosphonopropanoic Acid
2-Thiocytidine
Pc(18:0/20:4_O)
Pc(Pgd2/P-18:0)
1-Nitropyrene
Pc(O-18:1/22:5_2O)
Pe(Pgf2Alpha/20:2(11Z,14Z))
Butyl (S)-3-Hydroxybutyrate Glucoside
Pe(Pgf2Alpha/22:2(13Z,16Z))
-log10(p-value)
log2(FC)
Log2(FC)
P-value
1.0
0.1
0.05
0.01
0.001

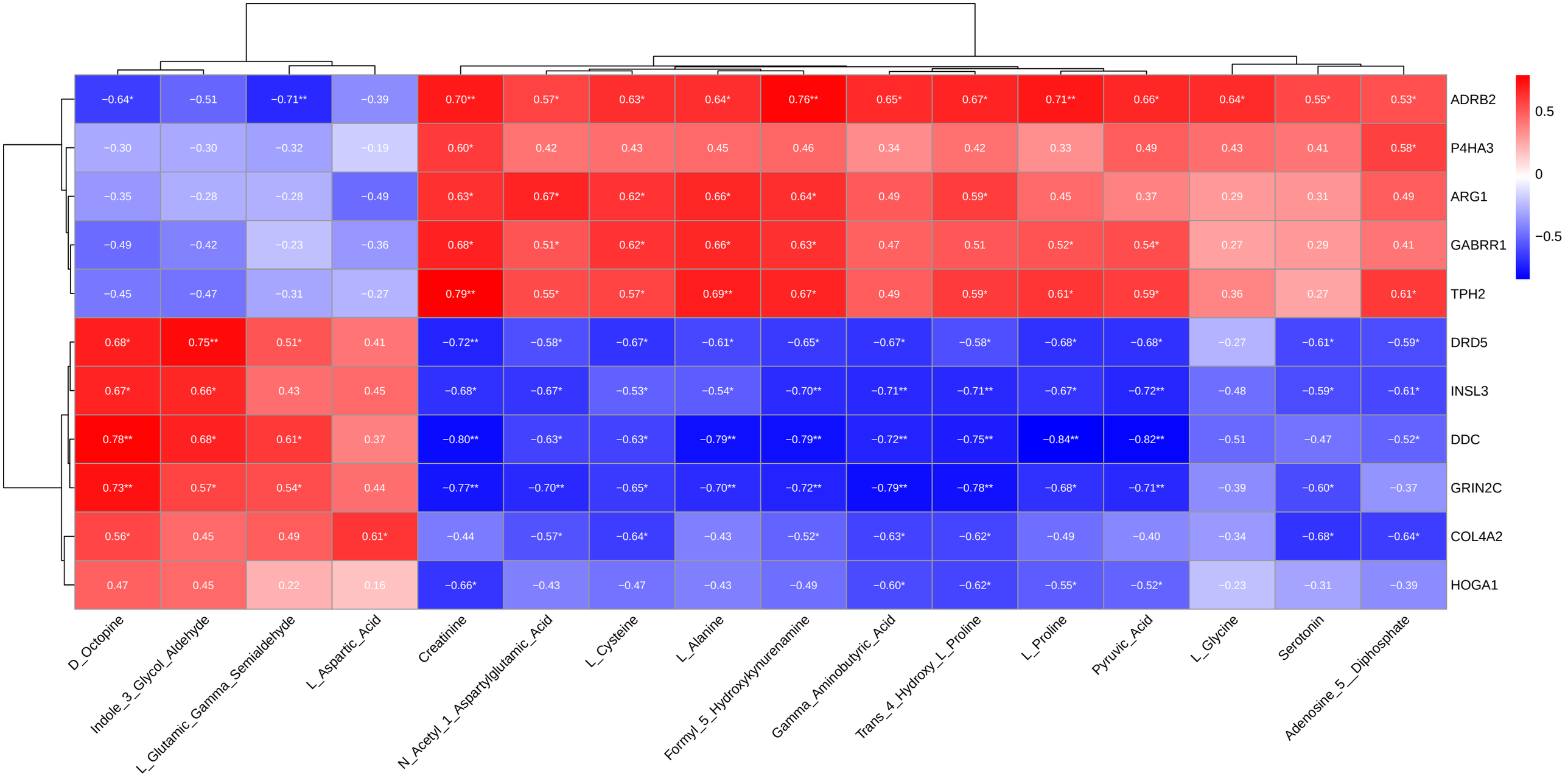
E.

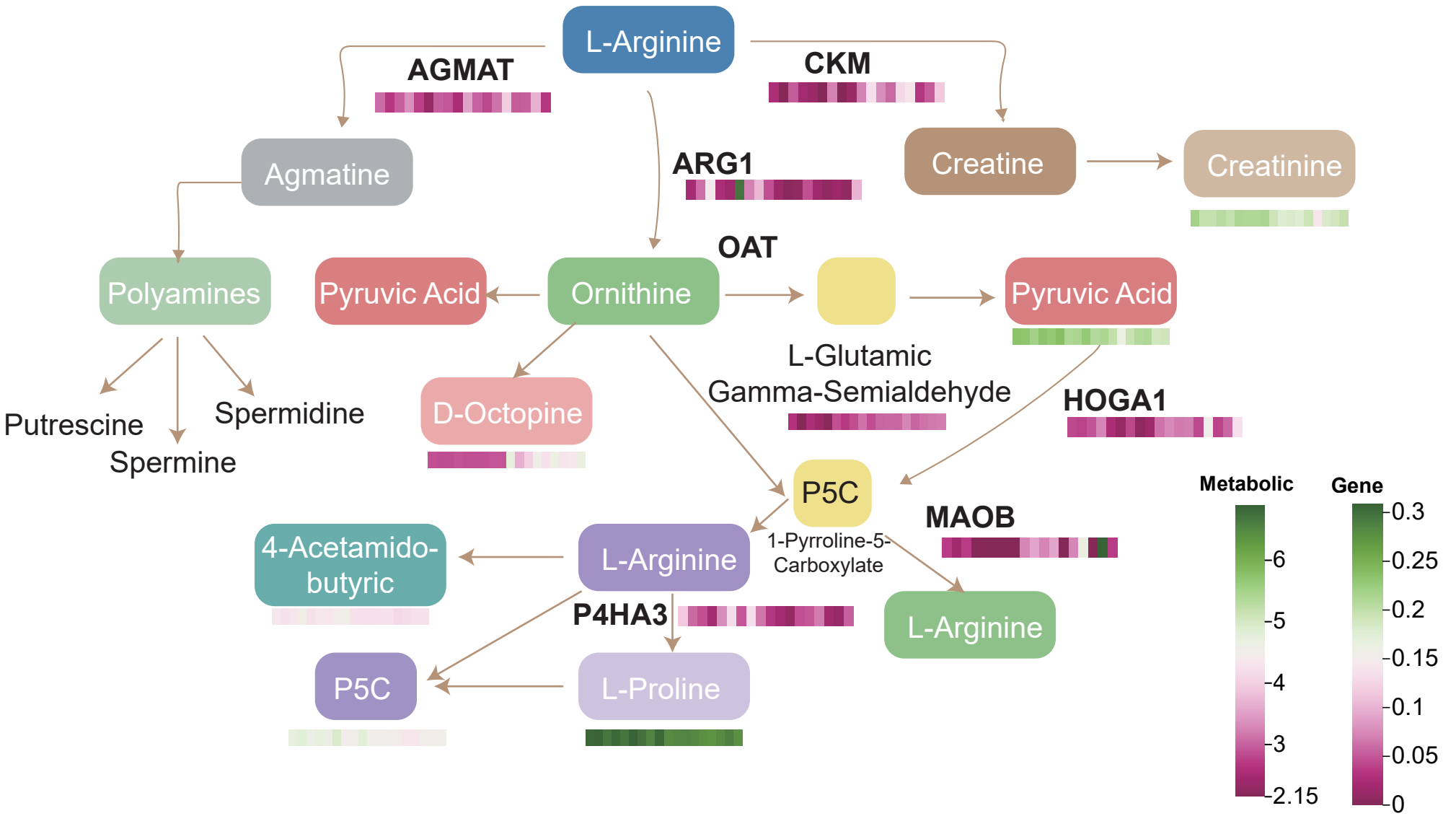

Insulin resistance
Aminoacyl-tRNA biosynthesis
Cysteine and methionine metabolism
Glycine, serine and threonine metabolism
Protein digestion and absorption
Synaptic vesicle cycle
Valine, leucine and isoleucine biosynthesis
Neuroactive ligand-receptor interaction
Primary bile acid biosynthesis
Pantothenate and CoA biosynthesis
Glutathione metabolism
Central carbon metabolism in cancer
Taurine and hypotaurine metabolism
Tryptophan metabolism
ABC transporters
Histidine metabolism
Arginine and proline metabolism
D-Amino acid metabolism
Bile secretion
Alanine, aspartate and glutamate metabolism
Rich Factor
P_value
Number


A.
PCA2 : 17.1%
PCA1 : 36.03%
HCG
BIG
B.
P2RY11
KLHL6
DPH2
DDC
ZBTB21
DUSP8
BNIP1
EFNA2
AKAP8
PDCD7
MEPCE
NFKBID
−log10 (P−value)
log2 (Fold Change)
sig
Down
Not
Up
C.
PLAU
NFKB2
HBEGF
OCLN
LTF
SLC4A1
CCR2
CXCL2
THBD
FGF18
PDGFB
VEGFA
NGF
EPOR
FOS
CX3CR1
CCL4
GPX3
HIF1A
FGF2
PLAUR
HSPA1A
POU5F1
OSM
CSF2
IL1B
ALB
IL4
FANCM
TNFAIP3
CXCL3
MUC1
REL
RELB
CXCR4
IL10
IL1A
KIT
WNT4
CDH2
NFKBIA
GRO1
EDN1
IL15
HBB
ICAM1
CCL19
CCL3
EGR1
ICOS
CXCR3
CXCL9
LIF
CSF3
IRF1
THBS1
F3
CD83
EPHA2
MRPS12
CCR8
NFKBIZ
ARG1
D.
Scale Free R²
Soft Threshold (power)
R²=0.8
R²=0.7
R²=0.6
E.
−0.74
(0.00047)
0.74
(0.00047)
−0.86
(5.4e−06)
0.86
(5.4e−06)
0.12
(0.63)
−0.12
(0.63)
0.86
(4.1e−06)
−0.86
(4.1e−06)
0.19
(0.46)
−0.19
(0.46)
0.59
(0.011)
−0.59
(0.011)
0.14
(0.59)
−0.14
(0.59)
HCG
BIG
F.
hydrogen peroxide catabolic process
cell adhesion
erythrocyte development
cell differentiation
oxygen transport
multicellular organism development
cell−matrix adhesion
inner ear morphogenesis
neuron differentiation
cell fate commitment
Biological Process
extracellular region
basement membrane
cell surface
hemoglobin complex
haptoglobin−hemoglobin complex
cytoskeleton
cytoplasm
external side of plasma membrane
glutamatergic synapse
cell projection
Cellular Component
calcium ion binding
peroxidase activity
heparin binding
haptoglobin binding
organic acid binding
protein binding
oxygen carrier activity
RNA polymerase II cis−regulatory region sequence−specific DNA binding
DNA−binding transcription factor activity, RNA polymerase II−specific
oxygen binding
Molecular Function
Cytoskeleton in muscle cells
ECM−receptor interaction
Pancreatic secretion
Parathyroid hormone synthesis, secretion and action
Insulin secretion
Integrin signaling
PI3K−Akt signaling pathway
Malaria
Focal adhesion
Cytokine−cytokine receptor interaction
KEGG Pathway
−log10(PValue)
Gene Count
Fold Enrichment


A.
PCA2 : 10.43%
PCA1 : 14.78%
HCG
BIG
B.
−log10 (P−value)
log2 (Fold Change)
sig
Down
Not
Up
C.
Scale Free R²
Soft Threshold (power)
R²=0.8
R²=0.7
R²=0.6
D.
−0.96
(3e−10)
0.96
(3e−10)
0.18
(0.48)
−0.18
(0.48)
0.52
(0.027)
−0.52
(0.027)
0.088
(0.73)
−0.088
(0.73)
HCG
BIG
E.
Diff
WGCNA
182
380
485
F.
Cornified envelope formation
Metabolic pathways
GABAergic synapse
Collecting duct acid secretion
Efferocytosis
Gastric acid secretion
Leukocyte transendothelial migration
Estrogen signaling pathway
Galactose metabolism
Fructose and mannose metabolism
Fold Enrichment
−log10(PValue)
Gene Count


A.
PCA2 : 16.47%
PCA1 : 26.18%
HCG
BIG
B.
−log10 (P−val)
log2 (Fold Change)
sig
Down
Not
Up
C.
PCA2 : 7.12%
PCA1 : 29.43%
HCG
BIG
D.
−log10 (P−value)
log2 (Fold Change)
sig
Down
Not
Up
E.
circ36279
SPON2
bta-miR-431
TNN
INSIG1
AEN
SHD
TBX2
bta-miR-154b
IL20RA
circ35151
F.
32
5
50


ADRB2 −0.64* −0.51 −0.71** −0.39 0.70** 0.57* 0.63* 0.64* 0.76** 0.65* 0.67* 0.71** 0.66* 0.64* 0.55* 0.53*
P4HA3 −0.30 −0.30 −0.32 −0.19 0.60* 0.42 0.43 0.45 0.46 0.34 0.42 0.33 0.49 0.43 0.41 0.58*
ARG1 −0.35 −0.28 −0.28 −0.49 0.63* 0.67* 0.62* 0.66* 0.64* 0.49 0.59* 0.45 0.37 0.29 0.31 0.49
GABRR1 −0.49 −0.42 −0.23 −0.36 0.68* 0.51* 0.62* 0.66* 0.63* 0.47 0.51 0.52* 0.54* 0.27 0.29 0.41
TPH2 −0.45 −0.47 −0.31 −0.27 0.79** 0.55* 0.57* 0.69** 0.67* 0.49 0.59* 0.61* 0.59* 0.36 0.27 0.61*
DRD5 0.68* 0.75** 0.51* 0.41 −0.72** −0.58* −0.67* −0.61* −0.65* −0.67* −0.58* −0.68* −0.68* −0.27 −0.61* −0.59*
INSL3 0.67* 0.66* 0.43 0.45 −0.68* −0.67* −0.53* −0.54* −0.70** −0.71** −0.71** −0.67* −0.72** −0.48 −0.59* −0.61*
DDC 0.78** 0.68* 0.61* 0.37 −0.80** −0.63* −0.63* −0.79** −0.79** −0.72** −0.75** −0.84** −0.82** −0.51 −0.47 −0.52*
GRIN2C 0.73** 0.57* 0.54* 0.44 −0.77** −0.70** −0.65* −0.70** −0.72** −0.79** −0.78** −0.68* −0.71** −0.39 −0.60* −0.37
COL4A2 0.56* 0.45 0.49 0.61* −0.44 −0.57* −0.64* −0.43 −0.52* −0.63* −0.62* −0.49 −0.40 −0.34 −0.68* −0.64*
HOGA1 0.47 0.45 0.22 0.16 −0.66* −0.43 −0.47 −0.43 −0.49 −0.60* −0.62* −0.55* −0.52* −0.23 −0.31 −0.39
D_Octopine
Indole_3_Glycol_Aldehyde
L_Glutamic_Gamma_Semialdehyde
L_Aspartic_Acid
Creatinine
N_Acetyl_1_Aspartylglutamic_Acid
L_Cysteine
L_Alanine
Formyl_5_Hydroxykynurenamine
Gamma_Aminobutyric_Acid
Trans_4_Hydroxy_L_Proline
L_Proline
Pyruvic_Acid
L_Glycine
Serotonin
Adenosine_5__Diphosphate
0.5
0
−0.5


L-Arginine
AGMAT
CKM
Agmatine
ARG1
Creatine
Creatinine
OAT
Polyamines
Pyruvic Acid
Ornithine
Pyruvic Acid
L-Glutamic
Gamma-Semialdehyde
HOGA1
Putrescine
Spermidine
Spermine
D-Octopine
P5C
1-Pyrroline-5-
Carboxylate
MAOB
4-Acetamido-
butyric
L-Arginine
P4HA3
L-Arginine
P5C
L-Proline
Metabolic
6
5
4
3
2.15
Gene
0.3
0.25
0.2
0.15
0.1
0.05
0