



Analysis of Upstream Regulators, Networks, and Pathways Associated With the Expression Patterns of Polycystic Ovary Syndrome Candidate Genes During Fetal Ovary Development

Rafiatu Azumah¹, Katja Hummitzsch¹, Monica D. Hartanti^{1,2}, Justin C. St. John¹, Richard A. Anderson³ and Raymond J. Rodgers^{1*}

¹Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, SA, Australia, ²Faculty of Medicine, Universitas Trisakti, Jakarta, Indonesia, ³MRC Centre for Reproductive Health, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom

OPEN ACCESS

Edited by:

Stephen J. Bush,
University of Oxford, United Kingdom

Reviewed by:

Avi Lerner,
Imperial College London,
United Kingdom
Jenny Visser,
Erasmus Medical Center, Netherlands

*Correspondence:

Raymond J. Rodgers
ray.rodgers@adelaide.edu.au

Specialty section:

This article was submitted to
Human and Medical Genomics,
a section of the journal
Frontiers in Genetics

Received: 21 August 2021

Accepted: 25 November 2021

Published: 07 February 2022

Citation:

Azumah R, Hummitzsch K,
Hartanti MD, St. John JC,
Anderson RA and Rodgers RJ (2022)
Analysis of Upstream Regulators,
Networks, and Pathways Associated
With the Expression Patterns of
Polycystic Ovary Syndrome Candidate
Genes During Fetal
Ovary Development.
Front. Genet. 12:762177.
doi: 10.3389/fgene.2021.762177

Polycystic Ovary Syndrome (PCOS) is a multifactorial syndrome with reproductive, endocrine, and metabolic symptoms, affecting about 10% women of reproductive age. Pathogenesis of the syndrome is poorly understood with genetic and fetal origins being the focus of the conundrum. Genetic predisposition of PCOS has been confirmed by candidate gene studies and Genome-Wide Association Studies (GWAS). Recently, the expression of PCOS candidate genes across gestation has been studied in human and bovine fetal ovaries. The current study sought to identify potential upstream regulators and mechanisms associated with PCOS candidate genes. Using RNA sequencing data of bovine fetal ovaries (62–276 days, $n = 19$), expression of PCOS candidate genes across gestation was analysed using Partek Flow. A supervised heatmap of the expression data of all 24,889 genes across gestation was generated. Most of the PCOS genes fell into one of four clusters according to their expression patterns. Some genes correlated negatively (early genes; *C8H9orf3*, *TOX3*, *FBN3*, *GATA4*, *HMG2*, and *DENND1A*) and others positively (late genes; *FDFT1*, *LHCGR*, *AMH*, *FSHR*, *ZBTB16*, and *PLGRKT*) with gestational age. Pathways associated with PCOS candidate genes and genes co-expressed with them were determined using Ingenuity pathway analysis (IPA) software as well as DAVID Bioinformatics Resources for KEGG pathway analysis and Gene Ontology databases. Genes expressed in the early cluster were mainly involved in mitochondrial function and oxidative phosphorylation and their upstream regulators included *PTEN*, *ESRRG/A* and *MYC*. Genes in the late cluster were involved in stromal expansion, cholesterol biosynthesis and steroidogenesis and their upstream regulators included *TGFB1/2/3*, *TNF*, *ERBB2/3*, *VEGF*, *INSIG1*, *POR*, and *IL25*. These findings provide insight into ovarian development of relevance to the origins of PCOS, and suggest that multiple aetiological pathways might exist for the development of PCOS.

Keywords: polycystic ovary syndrome, mitochondrial dysfunction, stromal expansion, steroidogenesis, upstream regulators, fetal ovary

INTRODUCTION

Polycystic Ovary syndrome (PCOS) is a debilitating syndrome with reproductive, endocrine and metabolic symptoms, affecting up to 10% women of reproductive age, with about 72% suffering infertility due to anovulation (Joham et al., 2015). Hyperandrogenism, oligo-/anovulation and the presence of polycystic ovary morphology are the cardinal features of the syndrome. However, hyperinsulinemia, type 2 diabetes mellitus, and obesity are also associated with PCOS [reviewed in detail by Teede et al. (2010)]. Despite the high prevalence of the syndrome, a comprehensive mechanism elucidating its pathophysiology is still lacking. Research efforts to define the pathophysiology of PCOS have increased the scientific conundrum surrounding its developmental origin(s) due to the diverging nature of research outcomes.

The multifactorial aetiology of PCOS has been evident by numerous genomic studies ranging from candidate gene approaches to Genome-Wide Association Studies (GWAS) – confirming its genetic origin. Candidate gene studies have identified androgen receptor (*AR*) (Schüring et al., 2012), fibrillin 3 (*FBN3*) (Hatzirodos et al., 2011), and anti-Müllerian hormone (*AMH*) (Tata et al., 2018) among others to be related to PCOS. GWAS have also identified 19 PCOS susceptibility loci among the Han Chinese, European and Korean ancestry [reviewed in detail by McAllister et al. (2015); Jones and Goodarzi (2016); Hiam et al. (2019)]. More so, the fetal origin of PCOS has been studied in human (Davies et al., 2012; Echiburú et al., 2020; Mills et al., 2020) and animal models including mice (Sullivan and Moenter 2004), rats (Wu et al., 2010), sheep (Birch et al., 2003), and monkeys (Dumesic et al., 1997). The possible fetal predisposition of the syndrome has also been studied further in bovine and human fetal ovaries using qRT-PCR technique (Hartanti et al., 2020; Liu et al., 2020).

Hartanti, et al. (2020), Liu, et al. (2020) showed that all PCOS candidate genes in/near loci identified by GWAS, except small ubiquitin like modifier 1 pseudogene 1 (*SUMO1P1*), and three additional candidate genes *AR*, *AMH* and transforming growth factor beta 1 induced transcript 1 (*TGFB111*) were expressed in human and bovine fetal ovaries. Three distinct expression patterns were observed; some genes were highly expressed during early stages, others during late stages whilst others were expressed throughout gestation (**Table 1**). Notably, the mRNA levels of genes within the early and late groups significantly correlated with each other as well as mean age (Hartanti, et al., 2020; Liu, et al., 2020). The expression of these PCOS genes during fetal development has provided further insight to the fetal origin of the syndrome; necessitating further studies to delineate its pathophysiology. Thus, the molecular mechanisms and upstream regulators associated with PCOS candidate genes during fetal development that could probably lead to the expression of various phenotypes (such as lean, obese, insulin resistant) during adulthood still remain elusive.

Therefore, there is a need to identify the upstream regulators and pathways that operate when PCOS candidate genes are expressed during normal ovary development. This will provide more insight on the possible perturbations during fetal ovary development that could lead to PCOS in adulthood and also help clarify some of the conundra surrounding the genetic and fetal origins of the disorder. Although studies to delineate the roles of PCOS candidate genes during fetal development are on-going, acquisition of human fetal ovary samples across gestation, especially from the third trimester, is a limitation. Considering the strong similarities between human and bovine ovaries in morphology and physiology, gestational length and the propensity for singleton pregnancies as well as the similarity in expression of PCOS candidate genes between human and

TABLE 1 | The three distinct expression patterns of PCOS candidate genes observed during fetal ovary development according to Hartanti, et al. (2020), Liu, et al. (2020); early, late and throughout gestation.

Early	Late	Throughout
Fibrillin 3 (<i>FBN3</i>)	Insulin receptor (<i>INSR</i>)	Thyroid adenoma associated (<i>THADA</i>)
GATA binding protein 4 (<i>GATA4</i>)	Follicle stimulating hormone receptor (<i>FSHR</i>)	Erb-B2 receptor tyrosine kinase 4 (<i>ERBB4</i>)
High mobility group AT-hook 2 (<i>HMGAT2</i>)	Plasminogen receptor with a C-terminal lysine (<i>PLGRKT</i>)	DNA repair protein (<i>RAD50</i>)
TOX high mobility group box family member 3 (<i>TOX3</i>)	Zinc finger and BTB domain containing 16 (<i>ZBTB16</i>)	Chromosome 9 open reading frame 3 (<i>C8H9orf3</i>)
DENN domain-containing 1A (<i>DENND1A</i>)	Interferon regulatory factor 1 (<i>IRF1</i>)	Yes associated protein 1 (<i>YAP1</i>)
Luteinising hormone/chorionic gonadotrophin receptor (<i>LHCGR</i>)	Transforming growth factor beta 1 induced transcript 1 (<i>TGFB111</i>)	Ras-related protein (<i>RAB5B</i>)
Follicle stimulating hormone beta subunit (<i>FSHB</i>)	Luteinizing hormone/choriogonadotropin receptor (<i>LHCGR</i>)	Sulphite oxidase (<i>SUOX</i>)
Erb-B2 receptor tyrosine kinase 3 (<i>ERBB3</i>)	Anti-müllerian hormone (<i>AMH</i>)	Ca ²⁺ /calmodulin-dependent protein kinase (<i>KRR1</i>)
–	Androgen receptor (<i>AR</i>)	ADP ribosylation factor like GTPase 14 effector protein (<i>ARL14EP</i>)
–	–	Farnesyl-diphosphate farnesyltransferase 1 (<i>FDFT1</i>)
–	–	Nei like DNA glycosylase 2 (<i>NEIL2</i>)
–	–	Microtubule associated protein RP/EB family member 1 (<i>MAPRE1</i>)

TABLE 2 | Characteristics of bovine fetal ovaries and RNA quality index (RQI).

Sample ID	Identical samples	Crown-rump length [cm]	Gestational age [days]	Gestational period	RQI
15/R12t	Y	7.7	62	Early I	9.0
15/R85t	Y	9	66	Early I	8.5
15/R86t	N	12	76	Early I	8.9
15/R74t	Y	14	82	Early I	9.1
15/R41t	N	17	91	Early II	9.3
15/R57t	N	19	98	Early II	9.2
15/R42t	Y	24	113	Early II	9.6
15/R51t	Y	28	124	Early II	9.6
15/R43t	Y	32	135	Early II	9.5
15/R2t	Y	39	154	Late	9.8
15/R44t	Y	45	170	Late	9.3
15/R1t	Y	58	201	Late	9.5
15/R45t	Y	74	234	Late	9.6
15/R33t	Y	80	245	Late	9.6
15/R47t	N	86	255	Late	9.4
15/R35t	Y	88	258	Late	9.5
15/R38t	Y	91	263	Late	9.3
15/R49t	N	93	266	Late	9.7
15/R50t	Y	100	276	Late	9.5

Identical samples refer to samples also analysed in Hartanti, et al. (2020), Liu, et al. (2020).

bovine during early stages of fetal ovary development (Hartanti, et al., 2020; Liu, et al., 2020), this study seeks to define the upstream regulators and pathways that operate when PCOS candidate genes are expressed (Hartanti, et al., 2020; Liu, et al., 2020) using bovine fetal ovaries.

MATERIALS AND METHODS

Bovine Fetal Ovary Collection

Fetal ovarian pairs across gestation (62–276 days, $n = 19$) were collected from pregnant *Bos taurus* cows at the abattoir of Midfield Meat International, Warrnambool, Victoria, Australia and were immediately frozen on dry ice on site and later stored in the laboratory at -80°C . These ovaries were scavenged from animals that were being processed for human consumption and were not owned by the authors or their institutions. As such the University of Adelaide's Animal Ethics Committee only requires notification of this. To estimate the gestational age of fetal samples, the crown-rump length (CRL) was measured (Russe 1983).

Sex Determination of Bovine Fetuses

Genomic DNA was extracted from the tail of fetuses with a CRL < 10 cm using the Wizard SV Genomic DNA Purification System (Promega Australia, Alexandria, NSW, Australia) according to the manufacturer's instructions. Two pairs of primers specific for a region in the Sex determining region Y (SRY) sequence (sense primer: 5'-TCACTCCTGCAAAAGGAGCA-3', antisense primer: 5'-TTATTGTGGCCAGGCTTG-3'), and for the 18S ribosomal RNA (18S) gene sequence were used for amplifying the genomic DNA in individual reactions as previously described (Hummitzsch et al., 2013).

RNA Extraction and RNA Sequencing

Whole fetal bovine ovaries were homogenised in 1 ml Trizol[®] (Thermo Fisher Scientific, Waltham, MA, United States) using the Mo Bio Powerlyser 24 (Mo Bio Laboratories Inc., Carlsbad, CA, United States) and RNA extracted according to manufacturer's instructions. All samples were treated with DNase I (Promega/Thermo Fisher Scientific Australia Pty Ltd., Tullmarine, Vic, Australia). The RNA concentration and quality (RQI, Table 2) was then determined using the Experion[™] RNA StdSens Analysis kit and the Experion[™] Automated Electrophoresis System (Bio-Rad Laboratories Pty., Ltd., Gladesville, NSW, Australia). 500 ng/50 μl per well (96-well plate) of total RNA of each sample was used for RNA-seq.

RNA-seq based transcriptome profiling was performed at the SAHMRI Genomics Facility (SAHMRI, Adelaide, SA, Australia). Briefly, Single-end Poly A-selection mRNA libraries (~35 M reads per sample) were created using the Nugen Universal Plus mRNA-Seq library kit from Tecan (Mannedorf, Switzerland) and sequenced with an Illumina Nextseq 500 using single read 75 bp (v2.0) sequencing chemistry (Illumina Inc., San Diego, CA, United States). Two sequencing runs, with 10 samples per run, were performed and sample 15/R43t was used as internal control in both runs.

RNA-Seq Data Analysis Using Partek Flow[®]

The raw data containing FASTQ files were uploaded to Partek Flow[®] Software, version 8.0 (Partek Incorporated, St. Louis, Missouri, United States). All samples underwent a pre-alignment quality assessment and showed Phred Quality Scores larger than 30. The reads were aligned and annotated to the bovine genome ARS-UCD1.2 (bosTau9; https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1/) using STAR 2.7.3a aligner ($>97\%$ alignment rate for all samples) and Partek

E/M, respectively. Transcript abundances were determined and expression levels presented as normalised counts per million (CPM). Initial comparison of gene expression profiles in the samples was then carried out using principal components analysis (PCA) (Hotelling 1933).

Analysis of PCOS Candidate Genes

Normalised RNA-seq data of bovine fetal ovaries across gestation were analysed to study the expression of PCOS candidate genes as well as their associated upstream regulators and mechanisms. The expression of 27 PCOS candidate genes, most of which are located in/near the loci associated with PCOS from previous GWAS, and three additional candidate genes *AR*, *AMH*, and *TGFBI1* were analysed using to Partek Flow® Software (version 8.0). Scatter plots showing the expression patterns for each candidate gene across gestation were also generated using GraphPad Prism version 8 (GraphPad Software Inc., La Jolla, Ca, United States). Pearson's correlation of the genes with each other as well as with gestational age were further analysed. Fold change and statistical significance of PCOS genes across gestation (comparing late versus early) were determined using the default parameters of Gene Specific Analysis (GSA); a multi-model approach based on Akaike Information Criterion (AIC) in Partek Flow®.

A supervised hierarchical clustering (heatmap) of all 24,889 genes identified in the RNA-seq data for all samples across gestation was then carried out; the location of PCOS genes on the heatmap was determined and clusters of PCOS genes as well as genes co-expressed with them were identified. Four clusters associated with the PCOS candidate genes were further studied using the core analysis component of Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, CA, United States). The canonical pathways, upstream regulators and networks associated with each cluster were analysed based on the statistical significance (*p*-value) and fold change (\log_2 ratio) of its gene list, determined using the default parameters of GSA in Partek Flow®. *C9orf3*, also known as aminopeptidase (*AOPEP*) replaced its bovine variant, *C8H9orf3*, in IPA analysis. For a more in-depth understanding of pathways associated with genes in each of the four clusters, the Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics Resources 6.8 (<https://david.ncifcrf.gov/home.jsp>, Frederick, MD 21702, United States) (Huang et al., 2009; Sherman and Lempicki 2009), was further used to analyse the pathway enrichment of genes for each cluster from Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases for bovine organisms.

Statistical Analyses

Expression of PCOS genes across gestation was plotted using GraphPad Prism version 8 (GraphPad Software Inc., La Jolla, CA, United States). PCA and hierarchical clustering were performed using the Euclidian algorithm for dissimilarity with average linkage. Pearson's correlation table and

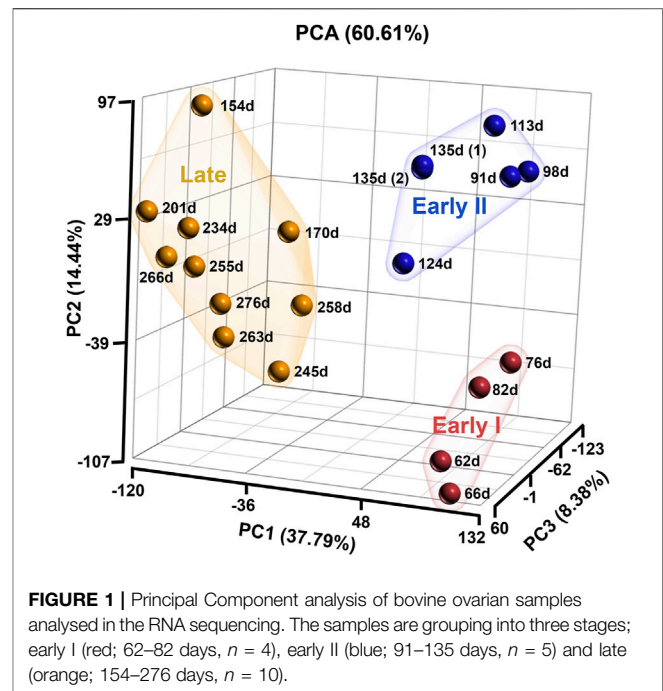


FIGURE 1 | Principal Component analysis of bovine ovarian samples analysed in the RNA sequencing. The samples are grouping into three stages; early I (red; 62–82 days, $n = 4$), early II (blue; 91–135 days, $n = 5$) and late (orange; 154–276 days, $n = 10$).

correlation coefficient (*R*) generated for scatterplots were determined in Partek Genomic Suite 7.0 (Partek Incorporated, St. Louis, Missouri, United States). The *p*-value of overlap between a set of significant molecules in each cluster and a given process/pathway/upstream regulator is determined using the Right-Tailed Fisher's Exact Test in IPA software. The *z*-score, which considers the directional effect of one molecule on another molecule/process and the direction of change of molecules in the dataset, represents activation when ≥ 2 and inhibition when ≤ -2 . R program, version 4.0.4 was used to plot GO graphs using data from DAVID bioinformatic resources.

RESULTS

Expression of PCOS Candidate Genes

The normalised RNA-seq data of bovine fetal ovaries identified 24,889 genes including uncharacterised and non-coding RNA. Principal Component Analysis (PCA) of all the genes in the data clustered the samples into three groups/stages; early I (62–82 days, $n = 4$), early II (91–135 days, $n = 5$), and late (154–276 days, $n = 10$) (Figure 1). For the purposes of this study, the two early groups were combined as “early” (62–135 days, $n = 9$) and all comparisons across gestation involved late versus early samples.

The expression of the 27 PCOS candidate genes identified in the RNA-seq data was studied across gestation. Using Pearson's correlation, the mRNA expression of these candidate genes across gestation showed three patterns; early, late, and throughout gestation. Thus, some PCOS candidate genes correlate significantly with gestational age negatively whilst others

positively (**Table 3**). Genes that correlate negatively with gestational age, and therefore named “early” genes, correlate positively with each other, whilst those that correlate positively with gestational age, named “late” genes, correlate positively with each other but negatively with the “early” genes. Thus, according to this study, the “early” genes consist of *C8H9orf3*, *ARL14EP*, *MAPRE1*, *TOX3*, *FBN3*, *GATA4*, *HMGA2*, and *DENND1A*; the “late” genes include *YAPI*, *INSR*, *THADA*, *TGFB11I*, *ZBTB16*, *IRF1*, *LHCGR*, *FSHR*, *AMH*, and *PLGRKT*; whilst genes expressed throughout gestation are *NEIL2*, *RAB5B*, *KRR1*, *SUOX*, *FDFT1*, *ERBB3*, *AR*, *ERBB4*, and *RAD50* (**Table 3**). Also, Gene Specific Analysis (GSA) comparing the expression of these candidate genes across gestation showed significant fold change across gestation; 14 of the genes were up-regulated and 13 down-regulated with significance at $p \leq 0.05$ for most candidate genes as shown in **Supplementary Table S1**. In addition to the “early” and “late” genes identified using Pearson’s correlation, other genes that correlate significantly with gestational age according to GSA are; *ERBB3* (negatively) as well as *AR* and *ERBB4* (positively); see **Supplementary Table S1**, **Supplementary Figure S1**.

A supervised hierarchical clustering of all 24,889 genes showed that most PCOS candidate genes formed close clusters with each

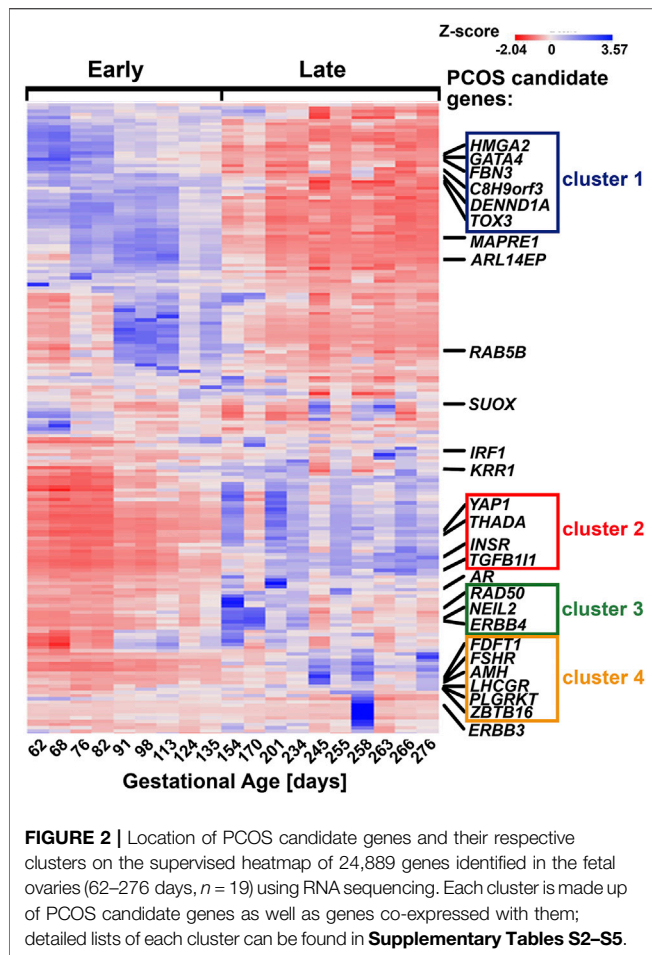
other according to the patterns earlier observed. Thus, most of the PCOS candidate genes that correlate significantly with each other in **Table 3** formed clusters closely with each other on the heatmap (**Figure 2**). Two strong clusters and two weak clusters were formed. The strong clusters contain “early” genes (cluster 1); *C8H9orf3*, *TOX3*, *FBN3*, *GATA4*, *HMGA2*, and *DENND1A* and “late” genes (cluster 4); *FDFT1*, *LHCGR*, *AMH*, *FSHR*, *ZBTB16* and *PLGRKT* (**Figure 2**). The two weak clusters comprised of “late” genes (cluster 2), *YAPI*, *INSR*, *THADA*, and *TGFB11I* and genes expressed throughout gestation (cluster 3), *RAD50*, *NEIL2*, and *ERBB4*. It is worth noting that although *FDFT1* does not correlate significantly with gestational age, it correlates significantly with most of the genes in cluster 4 according to the Pearson’s correlation (**Table 3**). Fold change and statistical significance were determined for the gene list of each of the clusters which were composed of PCOS candidate genes as well as the genes co-expressed with them (**Supplementary Tables S2–S5**). They were mapped to the IPA knowledge base and DAVID bioinformatics database to identify the canonical pathways, upstream regulators and networks associated with each cluster.

The findings of this study were then aligned with the developmental stages of the ovary. Both IPA and GO

TABLE 3 | Pearson’s correlation coefficients (R) between PCOS candidate genes mRNA expression levels and gestational age (n = 19) in bovine fetal ovaries.

Variable	Age (days)	HMGA2	TOX3	FBN3	GATA4	MAPRE1	DENND1A	ARL14EP	C8H9orf3	NEIL2	SUOX	ERBB3	RAB5B	KRR1	RAD50	FDFT1	ERBB4	AR	LHCGR	IRF1	AMH	PLGRKT	ZBTB16	FSHR	YAPI	THADA	INSR		
HMGA2	-0.946 ^a																												
TOX3	-0.93 ^b	-0.908 ^a																											
FBN3	-0.904 ^a	-0.933 ^b	0.826 ^b																										
GATA4	-0.903 ^a	-0.897 ^a	0.804 ^a	0.977 ^b																									
MAPRE1	-0.859 ^a	0.76 ^c	0.752 ^c	0.669 ^a	0.7 ^c																								
DENND1A	-0.728 ^a	0.747 ^b	0.744 ^b	0.74 ^b	0.728 ^b	0.452																							
ARL14EP	-0.716 ^b	0.547 ^b	0.664 ^b	0.397	0.444	0.869 ^b	0.298																						
C8H9orf3	-0.564 ^a	0.55 ^a	0.455	0.526 ^a	0.571 ^a	0.823 ^b	0.169	0.599 ^b																					
NEIL2	-0.431	0.328	0.441	0.198	0.2	0.331	0.596 ^b	0.448	-0.049																				
SUOX	-0.253	-0.182	0.038	0.397	0.48 ^a	0.087	0.207	-0.031	0.072	-0.238																			
ERBB3	-0.250	0.262	0.143	0.303	0.312	0.523 ^a	0.089	0.263	0.745 ^b	-0.236	0.164																		
RAB5B	-0.109	-0.07	-0.173	-0.211	-0.199	-0.117	0.085	0.536 ^a	-0.265	0.518 ^a	-0.171	-0.333																	
KRR1	-0.081	0.228	0.292	0.084	0.001	-0.046	0.243	0.03	-0.089	0.097	-0.314	-0.081	0.237																
RAD50	0.055	-0.034	0.07	-0.208	-0.256	-0.186	0.23	-0.037	-0.461 ^a	0.701 ^b	-0.49 ^a	-0.492 ^a	0.349	0.393															
FDFT1	0.222	-0.233	-0.4	-0.058	-0.059	-0.159	-0.192	-0.412	-0.037	-0.271	0.261	0.238	-0.512 ^a	-0.678 ^b	-0.402														
ERBB4	0.377	-0.368	-0.273	-0.486 ^b	-0.531 ^a	-0.453	0.017	-0.289	-0.635 ^b	0.537 ^a	-0.501 ^a	-0.463 ^a	0.26	0.123	0.835 ^b	-0.076													
AR	0.446	-0.309	-0.364	-0.412	-0.464 ^a	-0.557 ^b	-0.195	-0.508 ^a	-0.603 ^b	-0.193	-0.126	-0.348	-0.112	0.288	0.361	-0.015	0.407												
LHCGR	0.526 ^a	-0.489 ^a	0.561 ^a	-0.412	-0.434	-0.263	-0.41	-0.458 ^a	0.004	-0.453	-0.13	0.438	-0.423	-0.252	-0.314	0.639 ^b	0.023	0.106											
IRF1	0.602 ^a	-0.639 ^b	0.466 ^b	0.651 ^b	-0.651 ^b	-0.623 ^b	-0.232	-0.353	-0.676 ^b	0.08	-0.327	-0.525 ^a	0.306	-0.161	0.262	0.072	0.516 ^a	0.198	0.185										
AMH	0.643 ^b	-0.605 ^b	0.684 ^b	0.516 ^a	-0.524 ^a	-0.408	-0.481 ^a	-0.571 ^a	-0.153	-0.472 ^a	-0.138	0.158	-0.409	-0.29	-0.235	0.627 ^b	0.05	0.185	0.896 ^b	0.352									
PLGRKT	0.673 ^b	-0.588 ^b	0.755 ^b	-0.454	-0.484 ^a	-0.568 ^a	-0.626 ^b	-0.661 ^b	-0.22	-0.642 ^b	0.115	0.167	-0.514 ^a	-0.351	-0.361	0.64 ^a	0.0003	0.254	0.694 ^a	0.275	0.701 ^a								
ZBTB16	0.674 ^a	-0.649 ^b	0.656 ^b	0.682 ^b	-0.7 ^c	-0.59 ^b	-0.547 ^a	-0.527 ^a	-0.532 ^a	-0.167	-0.293	-0.455	-0.113	-0.427	0.094	0.5 ^a	0.384	0.38	0.425	0.614 ^a	0.597 ^b	0.487 ^a							
FSHR	0.704 ^a	-0.666 ^b	0.725 ^b	0.588 ^b	-0.602 ^b	-0.48 ^a	-0.463 ^a	-0.605 ^b	-0.243	-0.4	-0.195	0.166	-0.347	-0.255	-0.141	0.604 ^b	0.2	0.245	0.932 ^a	0.418	0.97 ^a	0.71 ^a	0.591 ^b						
YAPI	0.71 ^a	-0.555 ^a	0.557 ^a	0.558 ^a	-0.619 ^b	-0.863 ^b	-0.313	-0.775 ^a	-0.748 ^a	-0.243	-0.219	-0.47 ^a	-0.117	0.435	0.435	-0.108	0.514 ^a	0.764 ^a	0.164	0.379	0.281	0.351	0.374	0.357					
THADA	0.783 ^b	-0.636 ^b	0.664 ^b	0.628 ^b	-0.701 ^a	-0.838 ^b	-0.361	-0.841 ^a	-0.702 ^a	-0.258	-0.263	-0.277	-0.245	0.198	0.37	0.226	0.568 ^a	0.721 ^a	0.47 ^a	0.472 ^a	0.566 ^a	0.578 ^b	0.56 ^a	0.64 ^a	0.895 ^b				
INSR	0.907 ^a	-0.816 ^b	0.866 ^b	0.777 ^b	-0.797 ^b	-0.811 ^a	-0.691 ^b	-0.807 ^a	-0.562 ^a	-0.494 ^a	-0.22	-0.261	-0.336	-0.182	0.023	0.459 ^a	0.342	0.569 ^a	0.628 ^b	0.492 ^a	0.756 ^a	0.734 ^a	0.819 ^a	0.783 ^a	0.698 ^a	0.84 ^a			
TGFB11I	0.952 ^a	-0.904 ^b	0.929 ^a	0.822 ^a	-0.811 ^a	-0.848 ^a	-0.711 ^a	-0.747 ^a	-0.559 ^a	-0.496 ^a	-0.064	-0.268	-0.191	-0.254	-0.093	0.401	0.251	0.402	0.507 ^a	0.59 ^a	0.656 ^b	0.739 ^a	0.762 ^a	0.678 ^a	0.618 ^a	0.729 ^a			

Positive and negative correlations are marked in pink and blue, respectively. The colour intensity corresponds with the strength of the correlation. p-values: a < 0.05; b < 0.01; c < 0.001; d < 0.0001.



(biological processes) analyses of each cluster revealed relevant and similar canonical pathways to be associated with genes co-expressed with PCOS candidate genes across gestation. The top canonical pathways associated with the “early” genes (cluster 1) are mainly involved in mitochondrial function (Figure 3) whilst “late” genes were involved in stromal development and expansion (cluster 2) and lipid biosynthesis/steroidogenesis (cluster 4) (Figures 4, 5). Cluster 3, which consists of genes expressed throughout gestation, seems to be associated with a range of different pathways including central nervous system regulation and signalling among others (Figure 6). These pathways were also observed to some extent in the GO and KEGG pathway analysis for each cluster; with the strongest similarities to IPA observed in clusters 1 and 4 for mitochondrial function and steroidogenesis, respectively (Supplementary Tables S6–S9). Also, the number of genes extracted from each of these clusters as well as those that map to the IPA knowledge base and DAVID bioinformatic database for each cluster are shown in **Supplementary Table S10**.

Top biological upstream regulators associated with the strong clusters 1 and 4 and the weak clusters 2 and 3 were identified using IPA and are shown in **Tables 4, 5**, respectively. Upstream regulators such as *TGF β* , *TNF*, angiotensin, *ESR1*, among others were common to the “late” clusters 2 and 4. Upstream regulators

such as *PTEN*, *HNF4A*, *ESRRA/G*, *PSEN1*, *MYC*, mitochondrial *LONP1*, and *TP53* among others were associated with cluster 1 whereas *HNF1A*, *GATA2*, *PSENEN*, *IL25*, *REST*, *OCT4-NANOG*, *APHIA*, and *AGRN* were associated with cluster 3. Chemical upstream regulators (including endogenous molecules) associated with each cluster was also studied. The list of the top chemical (including endogenous molecules) upstream regulators associated with each cluster are detailed in **Supplementary Tables S11, S12**. Estradiol is among the top chemical upstream regulators associated with certain genes in clusters 1, 2, and 4 while dexamethasone regulates some genes in cluster 2 and 4 among others.

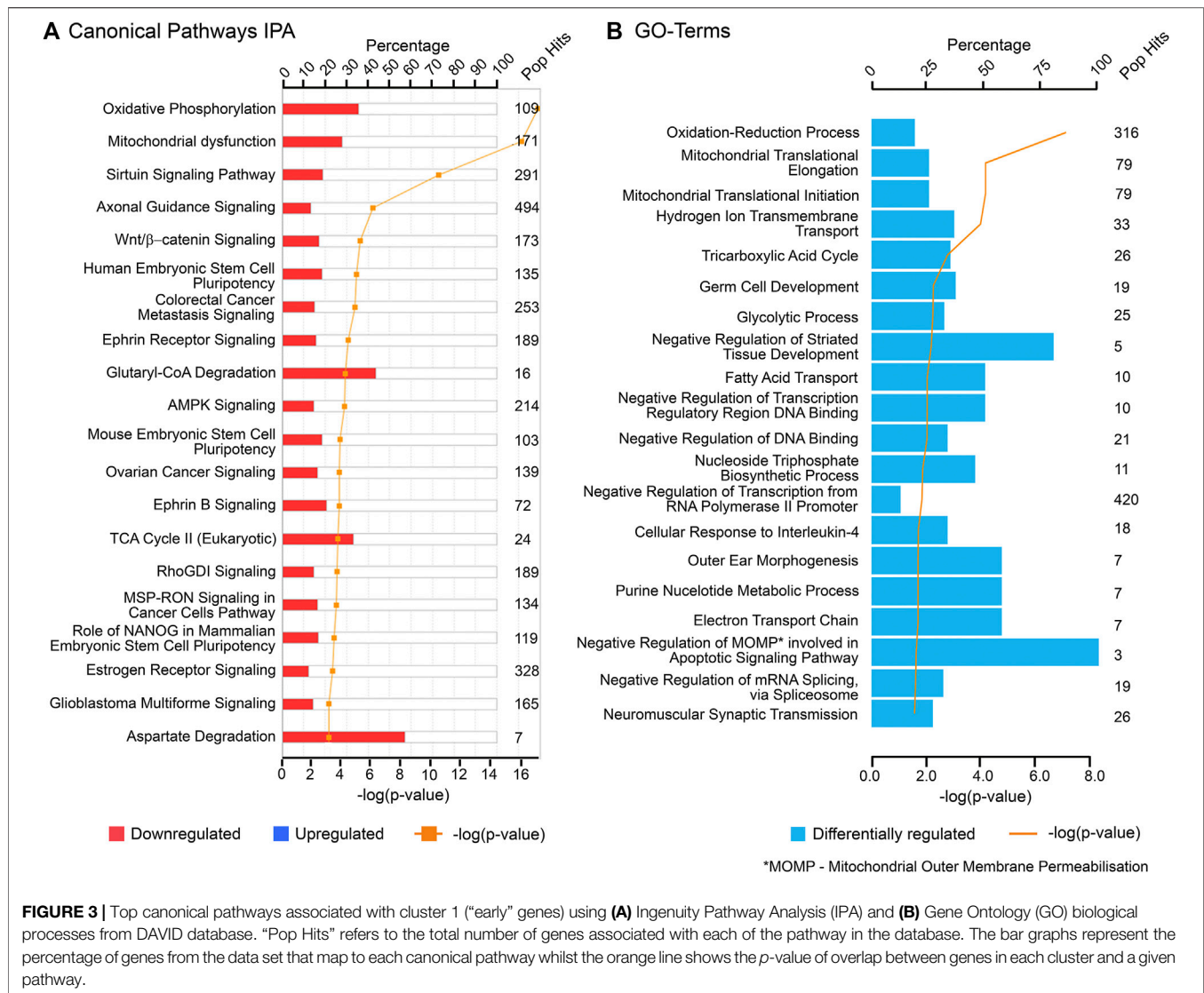
Additionally, 25 networks were generated by the IPA database for each cluster showing their functions and diseases. However, only those rated as important for fetal development and the role of PCOS candidate genes are shown in **Figures 7–9**. Specifically, mitochondrial function networks (1 and 3) of cluster 1 consisting of nuclear-encoded and mitochondrial DNA-encoded genes were downregulated during the second half of gestation (Figure 7). Networks associated with cluster 2, which were upregulated during the second half of gestation, consist of β -catenin as the central player (network 16) interacting with other subunits of its kind, namely frizzled molecules and cadherins, which are cell adhesion molecules (CAM) (Figure 8A). The second network of cluster 2 (network 24) consists of extracellular matrix components including the different types of collagen with two central molecules, collagen, and fibronectin (Figure 8B). The top networks for cluster 4, which were also upregulated during the second half of gestation, are associated with the effects of the transcription factor *MYC* in metabolism (network 1) and steroidogenesis (network 2) (Figure 9). All other relevant networks are summarized in **Supplementary Figures S2–S5**.

Nuclear-Encoded Mitochondrial Genes and Mitochondrial-DNA Encoded Genes

Oxidative phosphorylation and mitochondrial dysfunction pathways are the top canonical pathways co-expressed with the PCOS candidate genes expressed during early gestation (cluster 1). mRNA expression of nuclear-encoded mitochondrial genes associated with these top pathways as well as genes encoded by mitochondrial DNA from the RNA-seq are shown as scatterplots to outline their expression patterns across gestation (Supplementary Figures S6, S7, respectively). Similar patterns of expression were observed for both; thus, these genes are highly expressed during the early stages of gestation with expression decreasing gradually until 200 days of gestation, then the expression is maintained or increased slightly at that level in the third trimester.

DISCUSSION

Using RNA-seq data from normal fetal ovaries collected across gestation, we examined the expression patterns of PCOS



candidate genes and found that most belonged to one of four gene clusters. The expression patterns of most of the PCOS candidate genes in the RNA-seq data were consistent with those previously identified by both Hartanti, et al. (2020) and Liu, et al. (2020). In addition to the PCOS candidate genes identified as expressed “early” or “late” in gestation in these previous studies, we found that *MAPRE1*, *ARL14EP*, and *C8H9orf3* were also “early” genes, whilst *YAPI* and *THADA* were “late” genes. The study then identified canonical pathways, biological and chemical upstream regulators and networks of genes that are co-expressed in each of the clusters. This knowledge could be vital for understanding the mechanisms and identifying upstream regulators for delineating the fetal origin of PCOS.

During the early stages of fetal ovary development, primordial germ cells settle in the developing ovary as oogonia and start to proliferate between the proliferating Gonadal Ridge Epithelial-Like (GREL) cells. Then stroma of the mesonephros, containing fibroblasts, fibres, and

capillaries, penetrates into the gonadal ridge. This results in the formation of the ovigerous cords containing the GREL cells and proliferating oogonia. The ovigerous cords are separated from the penetrating stroma by a basal lamina and are open at the surface of the ovary during early gestation (Hummitzsch, et al., 2013; Hummitzsch et al., 2019). Interestingly, we found in cluster 1 that mitochondrial function is a key canonical pathway at this early stage of fetal ovary development. Expression patterns of nuclear-encoded mitochondrial genes and mitochondrial DNA-encoded genes across gestation showed that mitochondrial function and oxidative phosphorylation are at peak during stromal proliferation and penetration, specifically from the time of ovigerous cord formation until the cords begin to breakdown and primordial follicle formation occurs during fetal ovary development.

The relationship between mitochondrial dysfunction and the pathogenesis of PCOS remains unclear, although studies

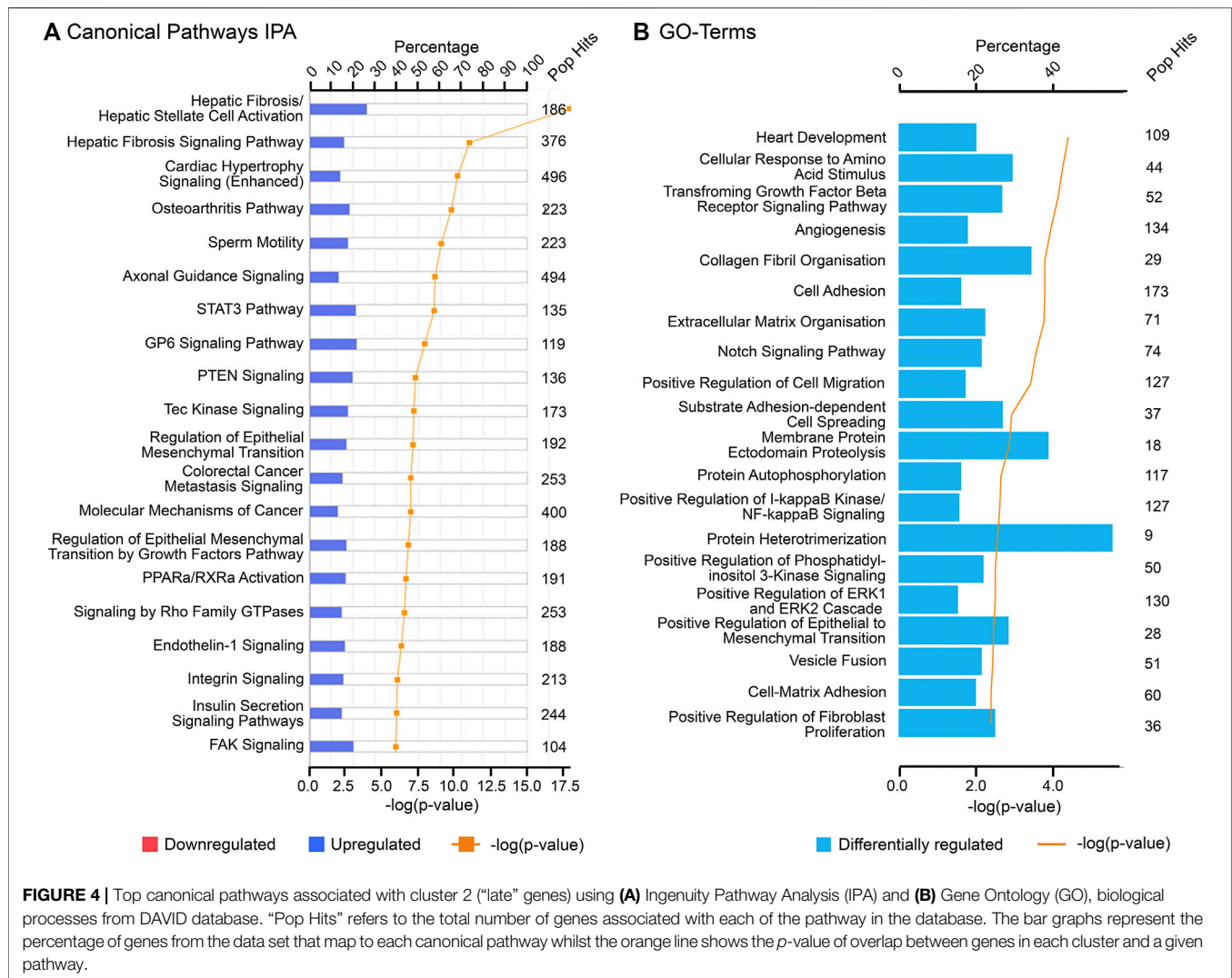
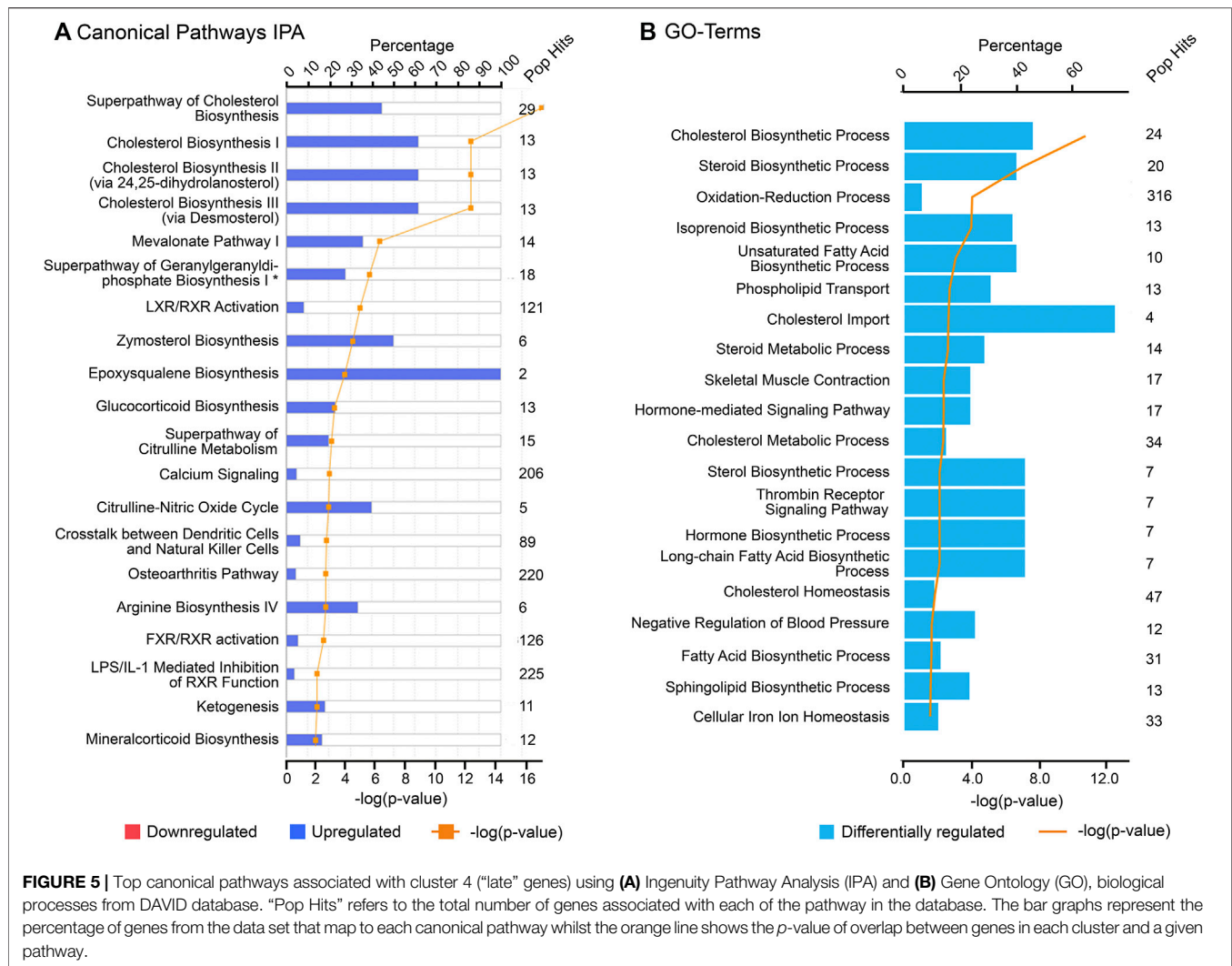


FIGURE 4 | Top canonical pathways associated with cluster 2 (“late” genes) using (A) Ingenuity Pathway Analysis (IPA) and (B) Gene Ontology (GO), biological processes from DAVID database. “Pop Hits” refers to the total number of genes associated with each of the pathway in the database. The bar graphs represent the percentage of genes from the data set that map to each canonical pathway whilst the orange line shows the p -value of overlap between genes in each cluster and a given pathway.

focussed on this topic began over the last decade. Mitochondrial DNA copy number and abnormal reactive oxidative species have been associated with numerous phenotypes of PCOS (Hyderali and Mala 2015; Shukla et al., 2020). Thus, insulin resistance, obesity, hyperandrogenism among others have been linked directly/indirectly to mitochondrial dysfunction [see reviews (Ilie 2018; Zhang et al., 2019; Zeng et al., 2020)]. A meta-analysis involving lean (BMI ≤ 23) and obese (BMI ≥ 23) PCOS patients using a systematic and comparative study by Idicula-Thomas et al. (2020) showed that nuclear-encoded mitochondrial genes were downregulated mainly in the cumulus cells of obese PCOS patients, but upregulated mainly in subcutaneous adipose tissues of lean PCOS patients. Furthermore, previous studies have shown that offspring of androgenised lean PCOS mice had impaired ovarian mitochondrial ultrastructure and function when

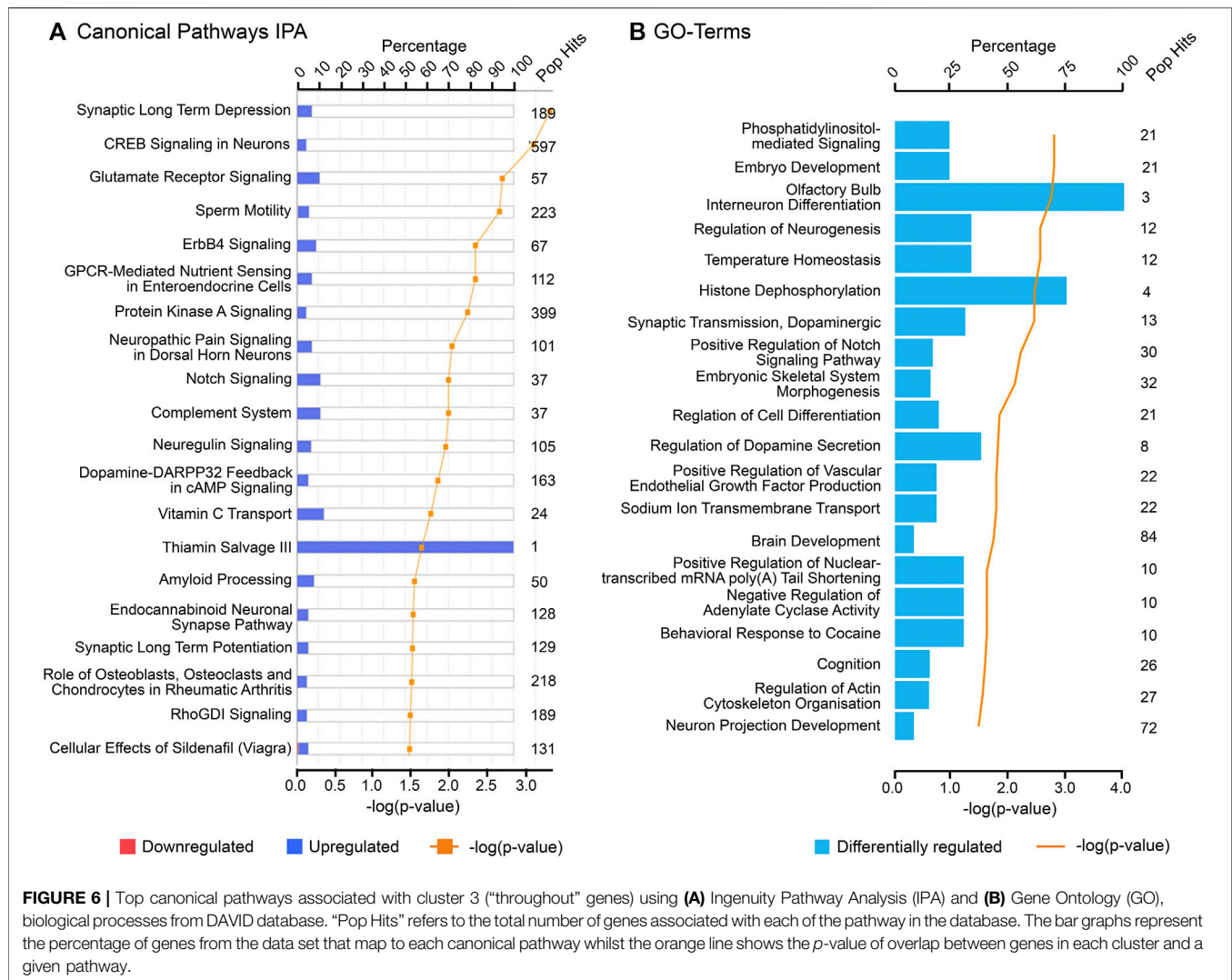
compared to their controls (Chappell et al., 2020). Thus, mitochondrial ultrastructure in oocytes of androgenised offspring had disorganised cristae and swollen vacuoles without any electron dense content; a lower inner mitochondrial membrane potential of oocytes was also observed as compromised mitochondria function in the PCOS lean mice (Chappell, et al., 2020). This implies that *in utero* exposures to high androgen levels can significantly affect the development of the ovaries, most probably at the early stages of gestation, when germ cells undergo mitosis, ovigerous cords form, and stroma proliferates and migrates leading to the fetal programming of PCOS. More so, aberrant mitochondrial function was observed in oocytes at germinal vesicles stage of PCOS patients when compared to healthy controls in a study that compared the various stages of oocyte development in the two groups using single cell RNA sequencing (Qi et al., 2020). Additionally,



studies focussed on the early stages of fetal ovary development in pig have also shown upregulation of mitochondrial activity and oxidative phosphorylation genes, inferred by higher expression of mitochondrial DNA-encoded genes at early stages of development (E31). During this stage, extensive proliferation of primordial germ cells and the global waves of methylation and demethylation occur in the porcine ovary; requiring significant levels of energy to sustain the changes (Zhu et al., 2021). The high energy produced results in by-products (e.g., alpha-ketoglutarate) that are then used to establish various epigenetic marks and modulate demethylation by oxidising of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (Tahiliani et al., 2009; Wu and Zhang 2017). The findings together infer that mitochondrial function could play a significant role in the predisposition of PCOS; requiring further studies.

Furthermore, the identification of upstream regulators such as *DAP3*, which plays significant roles in mitochondrial respiration and apoptosis (Tang et al., 2009), and *MYC*, which is a regulator of mitochondrial biogenesis (Li et al., 2005), as well as the downregulation of nuclear and mitochondrial DNA-encoded genes during the second half of gestation, in this study, implies that mitochondrial function is vital to the early stages of ovary development.

During mid-gestation, compartmentalisation of the ovary into the cortex and medulla becomes apparent, where there are alternating ovigerous cords and stromal areas in the cortex and stroma from the mesonephros containing extracellular matrix, fibroblasts and vasculature in the medulla. The stroma starts to migrate laterally after reaching just below the ovarian surface resulting in the closing of the ovigerous cords on the surface and the establishment of an ovarian surface epithelium underlain by a basal lamina (Hummitzsch, et al.,



2019; Hummitzsch, et al., 2013). This is, to some extent, consistent with our observations of genes connected with cluster 2, which includes genes involved in stromal expansion. Notably, the presence of elevated levels of fibrous tissues and collagen in the ovarian capsule or tunica albuginea resulting in denser ovarian stroma are common features of polycystic ovaries. Increased stromal collagen and ovarian cortex expansion in PCOS patients are associated with dysfunction of fibrillin 3 (*FBN3*) which is highly expressed in early stages of fetal development and not expressed late in gestation or adult ovaries (Hatzirodos, et al., 2011). *FBN3* regulates transforming growth factor- β (TGF- β) pathways, which stimulate fibroblast proliferation and collagen formation. TGF- β may also play significant roles in the cardiovascular and metabolic symptoms of PCOS as discussed in review by Raja-Khan et al. (2014). Upstream regulators involved in TGF- β signalling pathways such as

TGFB1, *TGFB2*, *TGFB3*, and *TGFB2* and fibroblast proliferation regulators such as *FGF2*, coagulation factor II (*F2*) were consistent with findings of this study. In addition, networks associated with the components of extracellular matrix, such as collagen and fibronectin which are relevant to this stage of ovary development were associated with this cluster of genes.

Furthermore, *YAP1*, a PCOS candidate gene in cluster 2, is a critical regulator of granulosa cell proliferation, differentiation and survival by interaction with epidermal growth factor receptor, gonadotrophin, and TGF- β signalling pathways (Lv et al., 2019). More so, *ERBB3*, which is an early gene according to our GSA analysis and Liu et al. (2020) and a member of the epidermal growth factor family, was identified as an upstream regulator for genes in cluster 2. Additionally, β -catenin, which is a central molecule for one of the networks associated with this cluster, plays an essential role in Wnt signal

TABLE 4 | Top biological upstream regulators and their respective activation z-score as well as *p*-value of association for the strong clusters, cluster 1 and 4.

	Upstream regulator	Name	z-score	p-value
Cluster 1	<i>CLPP</i>	Caseinolytic mitochondrial matrix peptidase proteolytic subunit	4.899	1.11E-11
	<i>PTEN</i>	Phosphatase and tensin homolog	-0.143	1.46E-10
	<i>HNF4A</i>	Hepatocyte nuclear factor 4 alpha	-3.972	1.90E-10
	<i>TLE3</i>	TLE family member 3, Transcriptional corepressor	—	2.69E-10
	<i>DAP3</i>	Death associated protein 3	-3.162	1.05E-09
	<i>ARNT</i>	Aryl hydrocarbon receptor nuclear translocator	-3.689	2.98E-09
	<i>LONP1</i>	Lon peptidase 1, Mitochondrial	-1.457	3.71E-09
	<i>ESRRG</i>	Estrogen related receptor gamma	-2.006	6.71E-09
	<i>Firre</i>	Firre intergenic repeating RNA element	-4.796	1.41E-08
	<i>APP</i>	Amyloid beta precursor protein	-2.107	3.01E-08
	<i>MYC</i>	MYC proto-oncogene, BHLH transcription factor	-7.632	3.11E-08
	<i>STK11</i>	Serine/Threonine kinase 11	-4.041	3.31E-08
	<i>ALKBH1</i>	AlkB homolog 1, Histone H2A dioxygenase	-2.646	8.62E-08
	<i>NSUN3</i>	NOP2/Sun RNA methyltransferase 3	-2.646	8.62E-08
	<i>PSEN1</i>	Presenilin 1	-0.649	1.64E-07
	<i>KDM5A</i>	Lysine demethylase 5A	4.914	2.45E-07
	<i>MAPT</i>	Microtubule associated protein tau	—	3.58E-07
	<i>RICTOR</i>	RPTOR independent companion of MTOR complex 2	4.854	5.14E-07
	<i>DDX5</i>	DEAD-box helicase 5	-3.638	2.46E-06
	<i>TP53</i>	Tumor protein P53	-1.511	2.90E-06
<i>Esrra</i>	Estrogen related receptor alpha	-1.849	3.18E-06	
Cluster 4	<i>SREBF2</i>	Sterol regulatory element binding transcription factor 2	4.586	1.41E-22
	<i>INSIG1</i>	Insulin induced gene 1	-4.447	5.74E-21
	<i>MAP2K5</i>	Mitogen-activated protein kinase 5	4.088	3.78E-18
	<i>SCAP</i>	SREBF chaperone	4.269	9.03E-18
	<i>POR</i>	Cytochrome P450 oxidoreductase	-3.776	4.93E-16
	<i>SREBF1</i>	Sterol regulatory element binding transcription factor 1	4.508	3.27E-14
	<i>MFSD2A</i>	Major facilitator superfamily domain containing 2A	-3.138	8.14E-13
	<i>ATP7B</i>	ATPase copper transporting beta	3.464	1.46E-12
	<i>MAPK7</i>	Mitogen-activated protein kinase 7	3.873	2.30E-12
	<i>NPPB</i>	Natriuretic peptide B	-3.283	3.41E-12
	<i>SH3TC2</i>	SH3 domain and tetratricopeptide repeats 2	3.162	2.27E-11
	<i>PROM1</i>	Prominin 1	-2.828	3.28E-11
	<i>PPARA</i>	Peroxisome proliferator activated receptor alpha	1.112	6.45E-11
	<i>NR5A1</i>	Nuclear receptor subfamily 5 group A member 1	3.446	7.22E-11
	<i>C4BP</i>	Complement component 4 binding protein	2.449	5.69E-10
	<i>RPTOR</i>	Regulatory associated protein of MTOR complex 1	3.2	1.43E-09
	<i>KIF3A</i>	Kinesin family member 3A	-2.596	1.47E-09
	<i>INSIG2</i>	Insulin induced gene 2	-2.586	1.97E-09
	<i>TGFB1</i>	Transforming growth factor beta 1	2.172	3.90E-09
	<i>SIRT2</i>	Sirtuin 2	1.913	6.77E-09

transduction and in intercellular adhesion by interacting with cadherin (Liu et al., 2009; Bush et al., 2019). Undoubtedly, the regulation of the PCOS candidate genes during this stage plays a substantial role in the cascade of events that occur later in fetal development.

During the late stages of ovary development, follicles consisting of oocytes and granulosa cells are formed from the differentiation of ovigerous cords (Hummitzsch, et al., 2013; Heeren et al., 2015). The first primordial follicles appear in the inner cortex-medulla region, surrounded by a basal lamina. The surface epithelium, mostly single-layered, then surrounds the ovary. Stroma beneath the surface epithelial basal lamina then develops into the tunica albuginea during the final developmental stages of the human and bovine ovary (Hummitzsch, et al., 2013; Heeren, et al., 2015; Hummitzsch, et al., 2019). Activation of some primordial follicles occurs leading to their

development into primary and preantral follicles. Remarkably, the genes in cluster 4 (including the PCOS candidate genes), top canonical pathways, upstream regulators and networks were involved in folliculogenesis and ovarian steroidogenesis which is consistent with occurrence during this stage of development. *FDFT1*, which is an essential enzyme in the synthesis of sterols further leading to the synthesis of cholesterol, correlates significantly with genes involved in folliculogenesis such as *AMH*, *FSHR*, *LHCGR*. More so, upstream regulators involved in steroid synthesis such as *SREBF2*, *INSIG1*, *TGFB1*, *RPTOR*, and networks such as network 2 which has a central molecule, Mitogen-Activated Protein Kinases, *ERK1/2*, interacting with other molecules of steroidogenesis such as FSH, LH, *AMH*, *CYP11A1* were also identified for this cluster.

Despite the increasing number of genetics studies to define the abnormalities associated with PCOS, the

TABLE 5 | Top biological upstream regulators and their respective activation z-score as well as *p*-value of association for the weak clusters, cluster 2 and cluster 3.

	Upstream regulator	Name	z-score	p-value	
Cluster 2	<i>TGFB1</i>	Transforming growth factor beta 1	7.017	1.77E-21	
	<i>HRAS</i>	HRas proto-oncogene, GTPase	-0.699	1.60E-14	
	<i>ERBB2</i>	Erb-B2 receptor tyrosine kinase 2	-0.667	7.95E-13	
	<i>F2</i>	Coagulation factor II, Thrombin	5.485	9.38E-13	
	Alpha catenin	Alpha catenin group	-5.256	1.76E-12	
	<i>TGFB2</i>	Transforming growth factor beta 2	3.062	3.71E-12	
	<i>TNF</i>	Tumor necrosis factor	3.597	1.00E-11	
	<i>Mek</i>	Mitogen-activated protein kinase 1	3.218	1.23E-11	
	<i>COLQ</i>	Collagen like tail subunit of asymmetric acetylcholinesterase	1.511	1.29E-11	
	<i>MRTFB</i>	Myocardin related transcription factor B	5.14	1.60E-11	
	<i>VEGF</i>	Vascular endothelial growth factor	5.581	9.26E-11	
	<i>TP53</i>	Tumor protein P53	4.493	9.90E-11	
	<i>Tgf beta</i>	Transforming growth factor beta	4.093	2.28E-10	
	<i>AGT</i>	Angiotensinogen	3.331	3.01E-10	
	<i>FGF2</i>	Fibroblast growth factor 2	4.223	4.14E-10	
	<i>ITGB1</i>	Integrin subunit beta 1	-1.223	5.19E-10	
	<i>SP1</i>	Sp1 transcription factor	5.601	1.01E-09	
	<i>ERBB3</i>	Erb-B2 receptor tyrosine kinase 3	0.325	1.71E-09	
	<i>TP63</i>	Tumor protein P63	1.448	1.80E-09	
	<i>TGFBR2</i>	Transforming growth factor beta receptor 2	3.079	3.79E-09	
	<i>TGFB3</i>	Transforming growth factor beta 3	3.759	4.62E-09	
	Cluster 3	<i>ASCL1</i>	Achaete-scute family BHLH transcription factor 1	2.543	7.44E-05
		<i>FOXQ1</i>	Forkhead box Q1	—	5.54E-04
<i>Dst</i>		Dystonin	—	5.54E-04	
<i>NOBOX</i>		NOBOX oogenesis homeobox	—	9.83E-04	
<i>HNF1A</i>		HNF1 homeobox A	2.88	1.51E-03	
<i>RSPO2</i>		R-spondin 2	—	3.22E-03	
<i>TBPL1</i>		TATA-box binding protein like 1	—	3.22E-03	
<i>ATOH1</i>		Atonal BHLH transcription factor 1	—	3.48E-03	
<i>GATA2</i>		GATA binding protein 2	-0.128	4.78E-03	
<i>SNCA</i>		Synuclein alpha	-1.673	4.83E-03	
<i>GRIN3A</i>		Glutamate ionotropic receptor NMDA type subunit 3A	-2.828	5.61E-03	
<i>SAMD4A</i>		Sterile alpha motif domain containing 4A	—	7.81E-03	
<i>PSENEN</i>		Presenilin enhancer, gamma-secretase subunit	—	1.08E-02	
<i>PTPN11</i>		Protein tyrosine phosphatase non-receptor type 11	—	1.09E-02	
<i>REST</i>		RE1 silencing transcription factor	-2.19	1.14E-02	
<i>OCT4-NANOG</i>		POU class 5 homeobox 1-nanog homeobox	—	1.41E-02	
<i>IL25</i>		Interleukin 25	0.277	1.63E-02	
<i>APH1A</i>		Aph-1 homolog A, gamma-secretase subunit	—	1.79E-02	
<i>AGRN</i>		Agtrin	—	1.79E-02	

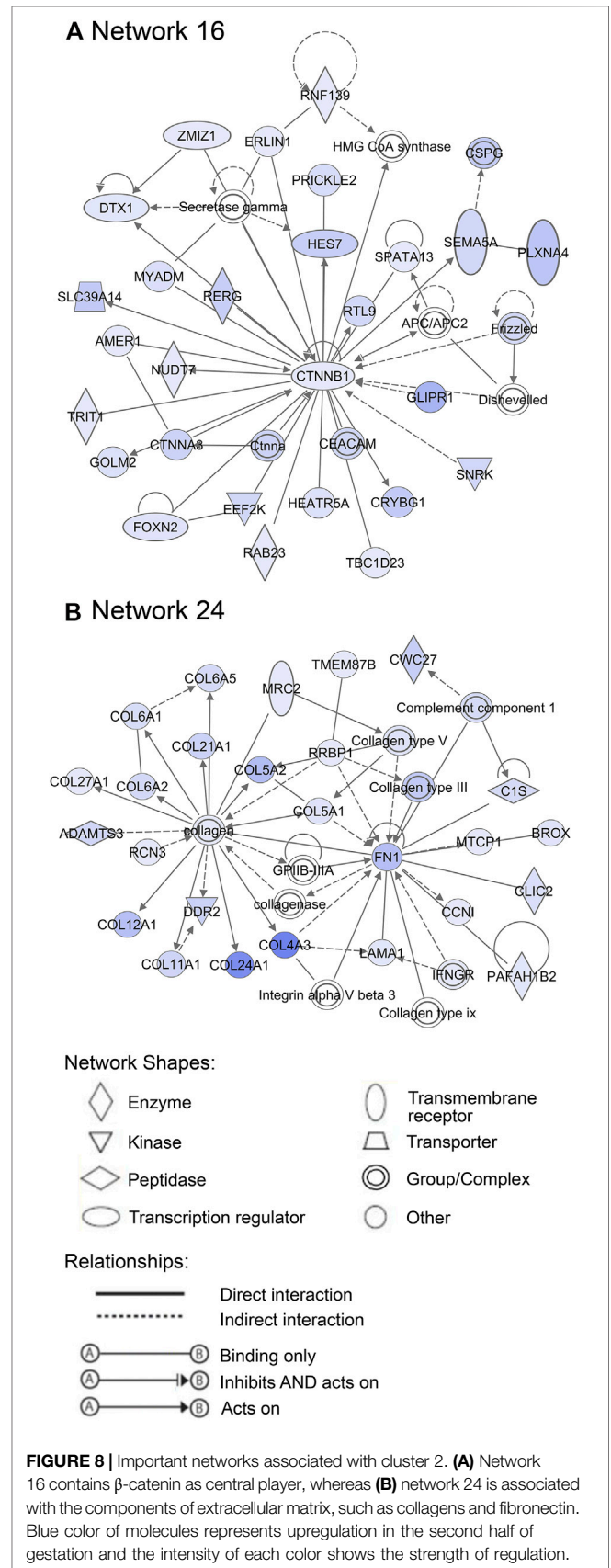
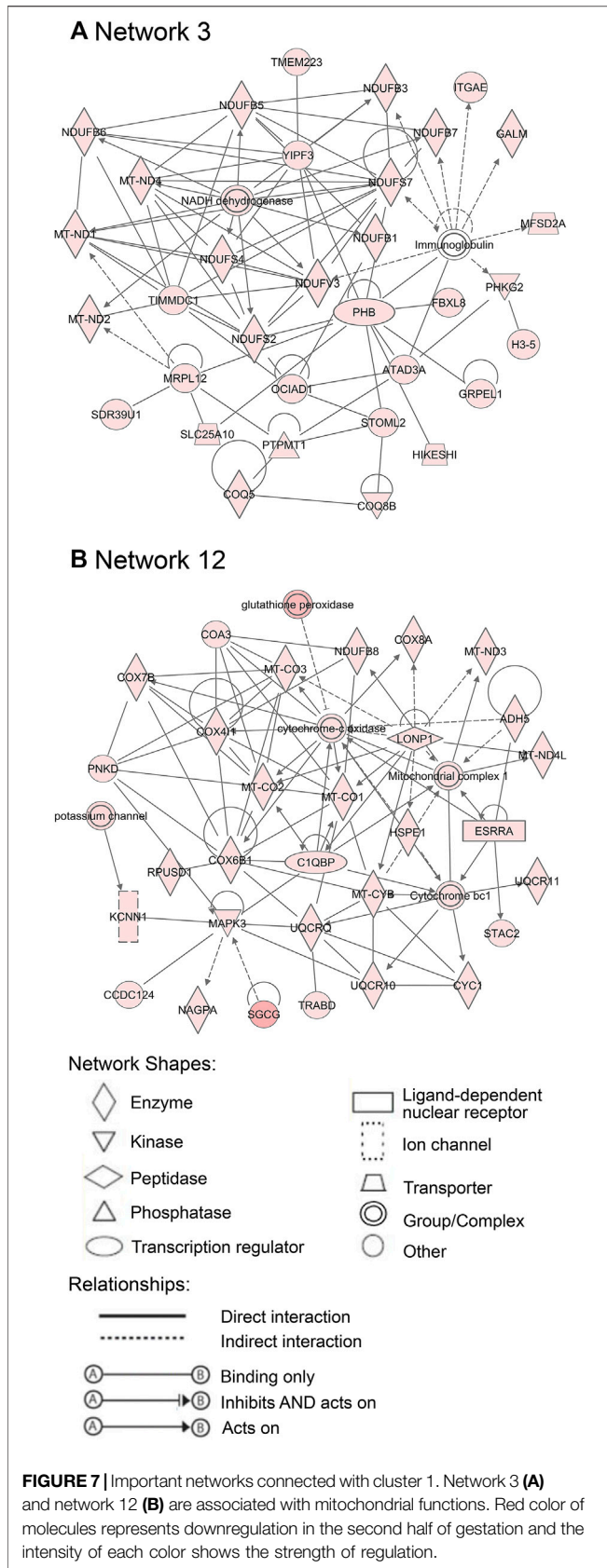
pathogenesis of the syndrome still remains a challenge. Numerous studies to delineate the transgenerational susceptibility of PCOS in animals and human are on-going (Tata, et al., 2018; Risal et al., 2019; Mimouni et al., 2021) and this study provides upstream regulators as well as canonical pathways that could be further studied to delineate the pathogenesis of the syndrome. Notably, genes co-expressed with most PCOS candidate genes during both the early and late developmental stages of fetal development are involved in pathways that have previously been associated with PCOS, although mechanisms remain elusive. It would be very logical to infer that perturbations associated with mitochondrial function in the early stages of fetal ovary development lead to a cascade of events across gestation, some of which impact key canonical pathways such as stromal expansion and steroidogenesis, possibly, leading to PCOS in adulthood. The outcome of this study supports the

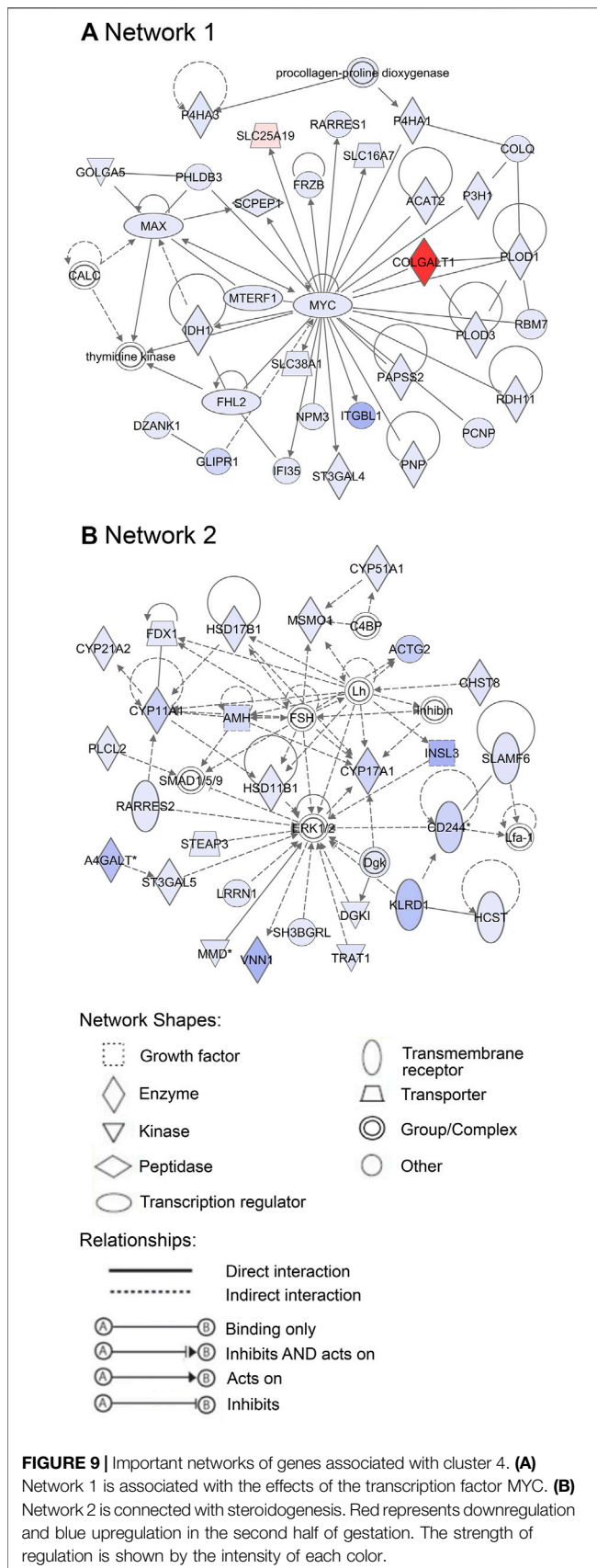
need to further study the fetal origin of PCOS with the hope to, not only define the syndrome, but also towards the diagnosis, treatment and prevention of syndrome.

In summary, although this study is an *in-silico* analysis of PCOS candidate genes and their co-expressed genes from RNA-seq data, it has shown significant consistency with the literature on ovarian development and PCOS. Additional studies to delineate the pathogenesis of the syndrome will be required based on these findings. The limitations of IPA software are also acknowledged as IPA is a knowledge based curated software.

CONCLUSION

These findings highlight the involvement of PCOS-associated genes in mitochondrial function, stromal expansion and





steroidogenesis during ovarian development. It could be speculated that perturbations during fetal ovary development resulting from dysregulation of these pathways could result in the different phenotypes of PCOS observed during adulthood. These perturbations could be genetic and/or environmental depending on maternal environment and fetal exposures associated with it. Further studies to delineate the role(s) of these pathways in the pathogenesis of PCOS are necessary. Notably, these findings infer a relationship between these major pathways, usually studied separately to define the syndrome and the need for studies investigating their interactions.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, GSE178450.

AUTHOR CONTRIBUTIONS

The authors contributions included Conceptualization (RA, KH, RJR), Formal analysis (RA, KH, RJR), Funding acquisition: KH, RJR, RAA, Investigation (RA), Analysis (RA, KH, RJR), Methodology (RA, KH, MDH, JCSJ, RJR), Project administration (KH, RJR), Resources (KH, RJR), Software (RA) Writing – original draft (RA) and Writing – review and editing (KH, JCSJ, RAA, RJR).

FUNDING

Funding was received from the Centre for Research Excellence on Women's Health in Reproductive life (GNT1171592), supported by the National Health and Medical Research Council of Australia and the UK Medical Research Council (G1100357).

ACKNOWLEDGMENTS

We thank Midfield Meat International, Warrnambool VIC, Australia for providing bovine fetal ovarian tissues. RA was supported by the Australian Research Training Program (fees-offset), the Faculty of Health and Medical Science Divisional Scholarship and the NHMRC CRE WHiRL PhD top-up scholarship. We are also grateful to the Partek team in Singapore for their quality support throughout the analysis process.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.762177/full#supplementary-material>

REFERENCES

- Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., et al. (2009). NCBI GEO: Archive for High-Throughput Functional Genomic Data. *Nucleic Acids Res.* 37, D885–D890. doi:10.1093/nar/gkn764
- Birch, R. A., Padmanabhan, V., Foster, D. L., Unsworth, W. P., and Robinson, J. E. (2003). Prenatal Programming of Reproductive Neuroendocrine Function: Fetal Androgen Exposure Produces Progressive Disruption of Reproductive Cycles in Sheep. *Endocrinology* 144, 1426–1434. doi:10.1210/en.2002-220965
- Bush, M., Alhanshali, B. M., Qian, S., Stanley, C. B., Heller, W. T., Matsui, T., et al. (2019). An Ensemble of Flexible Conformations Underlies Mechanotransduction by the Cadherin-Catenin Adhesion Complex. *Proc. Natl. Acad. Sci. USA* 116, 21545–21555. doi:10.1073/pnas.1911489116
- Chappell, N. R., Zhou, B., Schutt, A. K., Gibbons, W. E., and Blesson, C. S. (2020). Prenatal Androgen Induced Lean PCOS Impairs Mitochondria and mRNA Profiles in Oocytes. *Endocr. connections* 9, 261–270. doi:10.1530/ec-19-0553
- Davies, M. J., March, W. A., Willson, K. J., Giles, L. C., and Moore, V. M. (2012). Birthweight and Thinness at Birth Independently Predict Symptoms of Polycystic Ovary Syndrome in Adulthood. *Hum. Reprod.* 27, 1475–1480. doi:10.1093/humrep/des027
- Dumesic, D. A., Abbott, D. H., Eisner, J. R., and Goy, R. W. (1997). Prenatal Exposure of Female Rhesus Monkeys to Testosterone Propionate Increases Serum Luteinizing Hormone Levels in Adulthood. *Fertil. Sterility* 67, 155–163. doi:10.1016/s0015-0282(97)81873-0
- Echiburú, B., Milagro, F., Crisosto, N., Pérez-Bravo, F., Flores, C., Arpón, A., et al. (2020). DNA Methylation in Promoter Regions of Genes Involved in the Reproductive and Metabolic Function of Children Born to Women with PCOS. *Epigenetics* 15 (11), 1178–1194. doi:10.1080/15592294.2020.1754674
- Hartanti, M. D., Rosario, R., Hummitzsch, K., Bastian, N. A., Hatzirodos, N., Bonner, W. M., et al. (2020). Could Perturbed Fetal Development of the Ovary Contribute to the Development of Polycystic Ovary Syndrome in Later Life? *PLoS One* 15, e0229351. doi:10.1371/journal.pone.0229351
- Hatzirodos, N., Bayne, R. A., Irving-Rodgers, H. F., Hummitzsch, K., Sabatier, L., Lee, S., et al. (2011). Linkage of Regulators of TGF- β Activity in the Fetal Ovary to Polycystic Ovary Syndrome. *FASEB J.* 25, 2256–2265. doi:10.1096/fj.11-181099
- Heeren, A. M., Van Iperen, L., Klootwijk, D. B., de Melo Bernardo, A., Roost, M. S., Gomes Fernandes, M. M., et al. (2015). Development of the Follicular Basement Membrane during Human Gametogenesis and Early Folliculogenesis. *BMC Dev. Biol.* 15, 4. doi:10.1186/s12861-015-0054-0
- Hiam, D., Moreno-Asso, A., Teede, H. J., Laven, J. S., Stepto, N. K., Moran, L. J., et al. (2019). The Genetics of Polycystic Ovary Syndrome: An Overview of Candidate Gene Systematic Reviews and Genome-wide Association Studies. *Jem* 8, 1606. doi:10.3390/jcm8101606
- Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* 24, 417–441. doi:10.1037/h0071325
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res.* 37, 1–13. doi:10.1093/nar/gkn923
- Hummitzsch, K., Hatzirodos, N., Irving-Rodgers, H. F., Hartanti, M. D., Perry, V. E. A., Anderson, R. A., et al. (2019). Morphometric Analyses and Gene Expression Related to Germ Cells, Gonadal ridge Epithelial-like Cells and Granulosa Cells during Development of the Bovine Fetal Ovary. *PLoS One* 14, e0214130. doi:10.1371/journal.pone.0214130
- Hummitzsch, K., Irving-Rodgers, H. F., Hatzirodos, N., Bonner, W., Sabatier, L., Reinhardt, D. P., et al. (2013). A New Model of Development of the Mammalian Ovary and Follicles. *PLoS One* 8, e55578. doi:10.1371/journal.pone.0055578
- Hyderali, B. N., and Mala, K. (2015). Oxidative Stress and Cardiovascular Complications in Polycystic Ovarian Syndrome. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 191, 15–22. doi:10.1016/j.ejogrb.2015.05.005
- Iidicula-Thomas, S., Gawde, U., Bhaye, S., Pokar, K., and Bader, G. D. (2020). Meta-analysis of Gene Expression Profiles of Lean and Obese PCOS to Identify Differentially Regulated Pathways and Risk of Comorbidities. *Comput. Struct. Biotechnol. J.* 18, 1735–1745. doi:10.1016/j.csbj.2020.06.023
- Ilie, I. R. (2018). Advances in PCOS Pathogenesis and Progression-Mitochondrial Mutations and Dysfunction. *Adv. Clin. Chem.*, 127–155. doi:10.1016/bs.acc.2018.05.003
- Joham, A. E., Teede, H. J., Ranasingha, S., Zoungas, S., and Boyle, J. (2015). Prevalence of Infertility and Use of Fertility Treatment in Women with Polycystic Ovary Syndrome: Data from a Large Community-Based Cohort Study. *J. Women's Health* 24, 299–307. doi:10.1089/jwh.2014.5000
- Jones, M. R., and Goodarzi, M. O. (2016). Genetic Determinants of Polycystic Ovary Syndrome: Progress and Future Directions. *Fertil. Sterility* 106, 25–32. doi:10.1016/j.fertnstert.2016.04.040
- Li, F., Wang, Y., Zeller, K. I., Potter, J. J., Wonsey, D. R., O'Donnell, K. A., et al. (2005). Myc Stimulates Nuclearly Encoded Mitochondrial Genes and Mitochondrial Biogenesis. *Mol. Cell. Biol.* 25, 6225–6234. doi:10.1128/mcb.25.14.6225-6234.2005
- Liu, C. F., Bingham, N., Parker, K., and Yao, H. H. (2009). Sex-specific Roles of Beta-Catenin in Mouse Gonadal Development. *Hum. Mol. Genet.* 18, 405–417. doi:10.1093/hmg/ddn362
- Liu, M., Hummitzsch, K., Hartanti, M. D., Rosario, R., Bastian, N. A., Hatzirodos, N., et al. (2020). Analysis of Expression of Candidate Genes for Polycystic Ovary Syndrome in Adult and Fetal Human and Fetal Bovine Ovaries†. *Biol. Reprod.* 103, 840–853. doi:10.1093/biolre/iaaa119
- Lv, X., He, C., Huang, C., Wang, H., Hua, G., Wang, Z., et al. (2019). Timely Expression and Activation of YAP1 in Granulosa Cells Is Essential for Ovarian Follicle Development. *FASEB J.* 33, 10049–10064. doi:10.1096/fj.201900179rr
- McAllister, J. M., Legro, R. S., Modi, B. P., and Strauss, J. F., III (2015). Functional Genomics of PCOS: from GWAS to Molecular Mechanisms. *Trends Endocrinol. Metab.* 26, 118–124. doi:10.1016/j.tem.2014.12.004
- Mills, G., Badeghiesh, A., Suarathana, E., Baghlaif, H., and Dahan, M. H. (2020). Associations between Polycystic Ovary Syndrome and Adverse Obstetric and Neonatal Outcomes: a Population Study of 9.1 Million Births. *Hum. Reprod.* doi:10.1093/humrep/deaa144
- Mimouni, N. E. H., Paiva, I., Barbotin, A.-L., Timzoura, F. E., Plassard, D., Le Gras, S., et al. (2021). Polycystic Ovary Syndrome Is Transmitted via a Transgenerational Epigenetic Process. *Cel. Metab.* doi:10.1016/j.cmet.2021.01.004
- Qi, L., Liu, B., Chen, X., Liu, Q., Li, W., Lv, B., et al. (2020). Single-Cell Transcriptomic Analysis Reveals Mitochondrial Dynamics in Oocytes of Patients with Polycystic Ovary Syndrome. *Front. Genet.* 11, 396. doi:10.3389/fgene.2020.00396
- Raja-Khan, N., Urbanek, M., Rodgers, R. J., and Legro, R. S. (2014). The Role of TGF- β in Polycystic Ovary Syndrome. *Reprod. Sci.* 21, 20–31. doi:10.1177/1933719113485294
- Risal, S., Pei, Y., Lu, H., Manti, M., Fornes, R., Pui, H.-P., et al. (2019). Prenatal Androgen Exposure and Transgenerational Susceptibility to Polycystic Ovary Syndrome. *Nat. Med.* 25, 1894–1904. doi:10.1038/s41591-019-0666-1
- Rüsse, I. (1983). Oogenesis in Cattle and Sheep. *Bibl. Anat.* 24, 77–92.
- Schüring, A. N., Welp, A., Gromoll, J., Zitzmann, M., Sonntag, B., Nieschlag, E., et al. (2012). Role of the CAG Repeat Polymorphism of the Androgen Receptor Gene in Polycystic Ovary Syndrome (PCOS). *Exp. Clin. Endocrinol. Diabetes* 120, 73–79. doi:10.1055/s-0031-1291343
- Sherman, B. T., and Lempicki, R. A. (2009). Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* 4, 44. doi:10.1038/nprot.2008.211
- Shukla, P., Mukherjee, S., and Patil, A. (2020). Identification of Variants in Mitochondrial D-Loop and OrL Region and Analysis of Mitochondrial DNA Copy Number in Women with Polycystic Ovary Syndrome. *DNA Cel. Biol.* 39, 1458–1466. doi:10.1089/dna.2019.5323
- Sullivan, S. D., and Moenter, S. M. (2004). Prenatal Androgens Alter GABAergic Drive to Gonadotropin-Releasing Hormone Neurons: Implications for a Common Fertility Disorder. *Proc. Natl. Acad. Sci.* 101, 7129–7134. doi:10.1073/pnas.0308058101
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* 324, 930–935. doi:10.1126/science.1170116
- Tang, T., Zheng, B., Chen, S.-h., Murphy, A. N., Kudlicka, K., Zhou, H., et al. (2009). hNOA1 Interacts with Complex I and DAP3 and Regulates Mitochondrial Respiration and Apoptosis. *J. Biol. Chem.* 284, 5414–5424. doi:10.1074/jbc.m807797200

- Tata, B., Mimouni, N. E. H., Barbotin, A.-L., Malone, S. A., Loyens, A., Pigny, P., et al. (2018). Elevated Prenatal Anti-müllerian Hormone Reprograms the Fetus and Induces Polycystic Ovary Syndrome in Adulthood. *Nat. Med.* 24, 834–846. doi:10.1038/s41591-018-0035-5
- Teede, H., Deeks, A., and Moran, L. (2010). Polycystic Ovary Syndrome: a Complex Condition with Psychological, Reproductive and Metabolic Manifestations that Impacts on Health across the Lifespan. *BMC Med.* 8, 41. doi:10.1186/1741-7015-8-41
- Wu, X. Y., Li, Z. L., Wu, C. Y., Liu, Y. M., Lin, H., Wang, S. H., et al. (2010). Endocrine Traits of Polycystic Ovary Syndrome in Prenatally Androgenized Female Sprague-Dawley Rats. *Endocr. J.* 57, 201–209. doi:10.1507/endocrj.k09e-205
- Wu, X., and Zhang, Y. (2017). TET-mediated Active DNA Demethylation: Mechanism, Function and beyond. *Nat. Rev. Genet.* 18, 517–534. doi:10.1038/nrg.2017.33
- Zeng, X., Huang, Q., Long, S., Zhong, Q., and Mo, Z. (2020). Mitochondrial Dysfunction in Polycystic Ovary Syndrome. *DNA Cell Biol.* 39 (8), 1401–1409. doi:10.1089/dna.2019.5172
- Zhang, J., Bao, Y., Zhou, X., and Zheng, L. (2019). Polycystic Ovary Syndrome and Mitochondrial Dysfunction. *Reprod. Biol. Endocrinol.* 17, 67–15. doi:10.1186/s12958-019-0509-4
- Zhu, Q., Sang, F., Withey, S., Tang, W., Dietmann, S., Klisch, D., et al. (2021). Specification and Epigenomic Resetting of the Pig Germline Exhibit Conservation with the Human Lineage. *Cel. Rep.* 34, 108735. doi:10.1016/j.celrep.2021.108735
- Conflict of Interest:** Author RAA reports consultancy work for Ferring, Merck, IBSA, Roche Diagnostics, NeRRRe Therapeutics and Sojournix Inc.
- The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Azumah, Hummitzsch, Hartanti, St. John, Anderson and Rodgers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.