*Biology of Reproduction,* 2020, 103(3), 487–496
doi:10.1093/biolre/ioaa056
Research Article
Advance Access Publication Date: 21 April 2020



Research Article

# Distinct follicular and luteal transcriptional profiles in engineered human ectocervical tissue dependent on menstrual cycle phase

**Kelly E. McKinnon[1], Spiro Getsios[2] and Teresa K. Woodruff[1]**

[1]Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA and [2]Department of Dermatology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

***Correspondence:** Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, 303 E. Superior Street, Lurie 10-250, Chicago, IL 60611, USA. Tel: +312-503-2503; Fax: +312-503-0219; E-mail: tkw@northwestern.edu

**†Grant Support**: This work was supported by National Institutes of Health Common Fund (NCATS, NICHD, NIEHS, OWHR) UH3TR001207 and National Cancer Institute T32 CA009560 grants.

Received 5 December 2018; Revised 24 April 2019; Editorial Decision 5 November 2019; Accepted 0 Month 0000

## Abstract

To investigate genomic pathways that may influence physiology and infectivity during the menstrual cycle, RNA sequence analysis was performed on patient-matched engineered ectocervical tissue after follicular and luteal phase (LP) hormone treatments. We developed distinct cellular, molecular, and biological profiles in ectocervical epithelium dependent on the menstrual cycle phase. Follicular phase hormones were associated with proliferation, transcription, and cell adhesion, while LP samples expressed genes involved in immune cell recruitment, inflammation, and protein modifications. Additionally, our analysis revealed mucins not previously reported in ectocervical tissue, which could play an important role in fertility and disease prevention. This study provides insight into the phenomenon of increased LP vulnerability to infection and identifies potential targets for future research.

## Summary sentence

Next-generation sequencing of patient-derived ectocervical tissue models reveals distinct biological, cellular and molecular profiles based on menstrual cycle phase.

**Key words:** ectocervix, female reproductive tract, hormone action, mucins, differentiation, 3D model.

## Introduction

The cervix responds to ovarian hormones throughout the menstrual cycle each month in order to optimize conditions for fertility, while maintaining a chemical and physical barrier against pathological infection. Recent evidence suggests that due to these somewhat opposing functions, there may be a critical window of time in which infection is more likely, as immune defenses are downregulated to increase chances of fertilization [1, 2]. Additionally, many genes that have been previously shown to be involved in HIV infection [3–5] may be upregulated during the luteal phase (LP), as Yildiz-Arslan et al. showed via microarray analysis of gene expression in endocervical tissue. As a primary entry point for HIV and many other sexually transmitted infections, it is important to understand how fluctuating steroid hormones may influence pathogenic infection, so that effective preventative and therapeutic treatments can be developed.

Previously we demonstrated that engineered cervical tissue phenocopied native patient tissue (PT) morphology, responded to

hormones, expressed properly localized differentiation-associated proteins and mucins, and remained viable up to 28 days in culture [6]. Out of the three models we developed, an engineered ectocervical tissue most closely represented in vivo biology, making it the best choice for RNA sequence analysis of genes expressed during different phases of the menstrual cycle. Here, we present distinct molecular, cellular and biological profiles for human ectocervix during the follicular and LPs of the menstrual cycle due to differential transcript-isoform and gene expression in response to ovarian hormones, providing critical insight into ovarian hormone action in ectocervical epithelium during homeostasis.

## Methods

### Ethics statement

Ectocervical tissue samples were collected with written consent from women undergoing hysterectomies at Northwestern University Prentice Women's Hospital (Chicago, IL), according to an Institutional Review Board-approved protocol.

### PT acquisition and ectocervical model generation

Ectocervical tissue was obtained from consented patients undergoing hysterectomy for benign uterine conditions. Patient indication and generation of ectocervical tissue models have been described in depth in our previous study (BOR_1). Briefly, epithelial cells were isolated using techniques previously established for human keratinocyte cultures [4]. Epithelial cells were removed from stroma with forceps after 24 h incubation at 4C with dispase. Cells were dissociated using 0.25% trypsin and plated on 25% confluent feeder layers of J2-3T3 cells treated with mitomycin C (Calbiochem) for 2–2.5 h to induce cell cycle arrest. Feeder cells were cultured in fibroblast media (FM) until co-culture (DMEM HG (Sigma) supplemented with 10% neonatal calf serum (Fisher), 10 ug/ml gentamicin (Sigma), 0.25 ug/ml Amphotericin B (Cellgro)). Epithelial co-cultures with feeder cells were cultured in ectocervical media (EM) with additional 5% FBS and 10 ng/ml EGF (50% DMEM HG with glutamine (Sigma), 50% DMEM:F-12 (Sigma), supplemented with: 10 ug/ml gentamicin (Sigma), 0.25 ug/ml amphoterecin B (Cellgro), 0.4 ug/ml hydrocortisone (Sigma), 10 ng/ml cholera toxin (Sigma), 5% FBS (Fisher), and 1% ectocervical cocktail mix: 180 uM adenine (Sigma), 5 ug/ml human recombinant insulin (Sigma), 5 ug/ml human apotransferrin (Sigma), 5 ug/ml triiothyronine, T3 (Sigma). Feeder cells were changed every 2–3 days by incubating at RT in EDTA, which selectively removes fibroblasts, as described previously [4]. Stromal fibroblasts were isolated through explant outgrowth, as described previously [4], and cultured in FM.

To generate a cervical stromal equivalent, methods used to generate engineered dermis from neonatal foreskin were adapted with changes [4]. Briefly, we seeded $3 \times 10^5$ primary fibroblasts and J2-3T3 feeder cells in a collagen hydrogel consisting of 4 mg/ml rat tail collagen, 10% reconstitution buffer, 10% 10× DMEM (Sigma), 15 µl/ml 0.5 N NaOH, and $ddH_2O$ to a final volume of 2 ml collagen/well. Collagen gels were formed on 30 mm cell culture inserts (EMD Millipore Co), incubated at 37 °C for 30 min, and then submerged with FM. After 24 h, $10^6$ epithelial cells were seeded on each stromal equivalent and submerged with EM + 5% EGF above and below the culture insert. Once a monolayer of epithelial cells developed (5–7 days), engineered cervical tissue (ECT) was cultured at an air-liquid interface 5–7 days before beginning hormone treatments.

### Exogenous hormone treatments

Estradiol (E2) and progesterone (P4) solutions were made in 100% EtOH and stored at −20 °C between uses. To model ovarian hormone signaling during the follicular phase (FP) and LP, four levels of physiological concentrations of E2 and P4 were cycled through media during a 14 days culture as described previously (BOR_1). Engineered tissues were exposed to increasing levels of E2 (0.1–1.0 nM) for FP-7 treatments (days 0–7), and rising levels of P4 (10–50 nM) combined with decreasing E2 (1–0.01 nM) for LP-14 treatments (days 8–14) or cultured without hormones 7–14 days (Figure 1A).

### Total RNA isolation

Flash-frozen PT weighing up to 100 mg was pulverized using a Biopulverizer (Biospec, Bartlesville, OK) and placed in 700 µl of Qiazol (Qiagen, Venlo, Netherlands). Engineered tissue was washed with PBS ($Ca^{++}$, $Mg^{+}$) before removing tissue from culture-insert with razor blade. Epithelium was removed from collagen hydrogel with forceps and placed in 700 µl of Qiazol. Tissue solutions were needle-sheared and further broken down in Qiashredder columns (Qiagen). RNA was isolated using RNEasy mini-kit (Qiagen), according to the manufacturer's instructions. Concentrations of RNA were determined using a nanodrop spectrophotometer, and RNA was stored in RNAse-free water at −80 °C until analysis.

### Next generation sequencing

RNA (2 µg/sample) was isolated from epithelial cells used in hormonal experiments on 3D engineered tissue epithelium using cells isolated from two patients in four different conditions each. Conditions analyzed for each patient included control-no hormones day 7 (C-7), FP hormones day 7 (FP-7), control-no hormones day 14 (C-14), and LP hormones day 14 (LP-14). RNA samples were submitted to the University of Chicago Genomics Facility—Knapp Center for Biomedical Discovery (KCBD) (Chicago, IL) for quality check by BioAnalyzer, generation of cDNA library, and high-throughput sequencing with Illumina HiSeq4000. Raw reads were provided by KCBD for downstream analysis.

### Differential gene expression analysis

To map raw reads to the genome, open-source software Tophat2 was used through a UNIX command shell, as previously described [7]. Reads for all eight samples and a reference genome (hg_19) were loaded into TopHat2 to identify genomic positions for each fragment, and StringTie (ST) [8] was used to assemble mapped data into transcripts and quantify fragments per kilobase per million (FPKM) for each transcript detected. The "gffcompare" function was used to determine the number of all known and novel transcripts, genes, introns, and exons in merged samples compared to reference genome and to evaluate precision and sensitivity of alignment. StringTie output files for each sample were loaded into a dataframe in the Ballgown statistical analysis package in R [9]. For primary analyses, data were filtered to include only genes in which the row variance between samples was greater than one. Genes of particular interest, such as mucins, were evaluated in the unfiltered data. A value of 1 was added to all FPKM values before log2 transformation to inspect distribution of FPKM values. The confounding patient variable was corrected for using the "stattest" function in Ballgown with argument "adjustvar = patient". After filtering, genes found to be differentially expressed between sample groups ($P < 0.05$ by Kruskal–Wallis test) were used in cluster analysis. Each sample

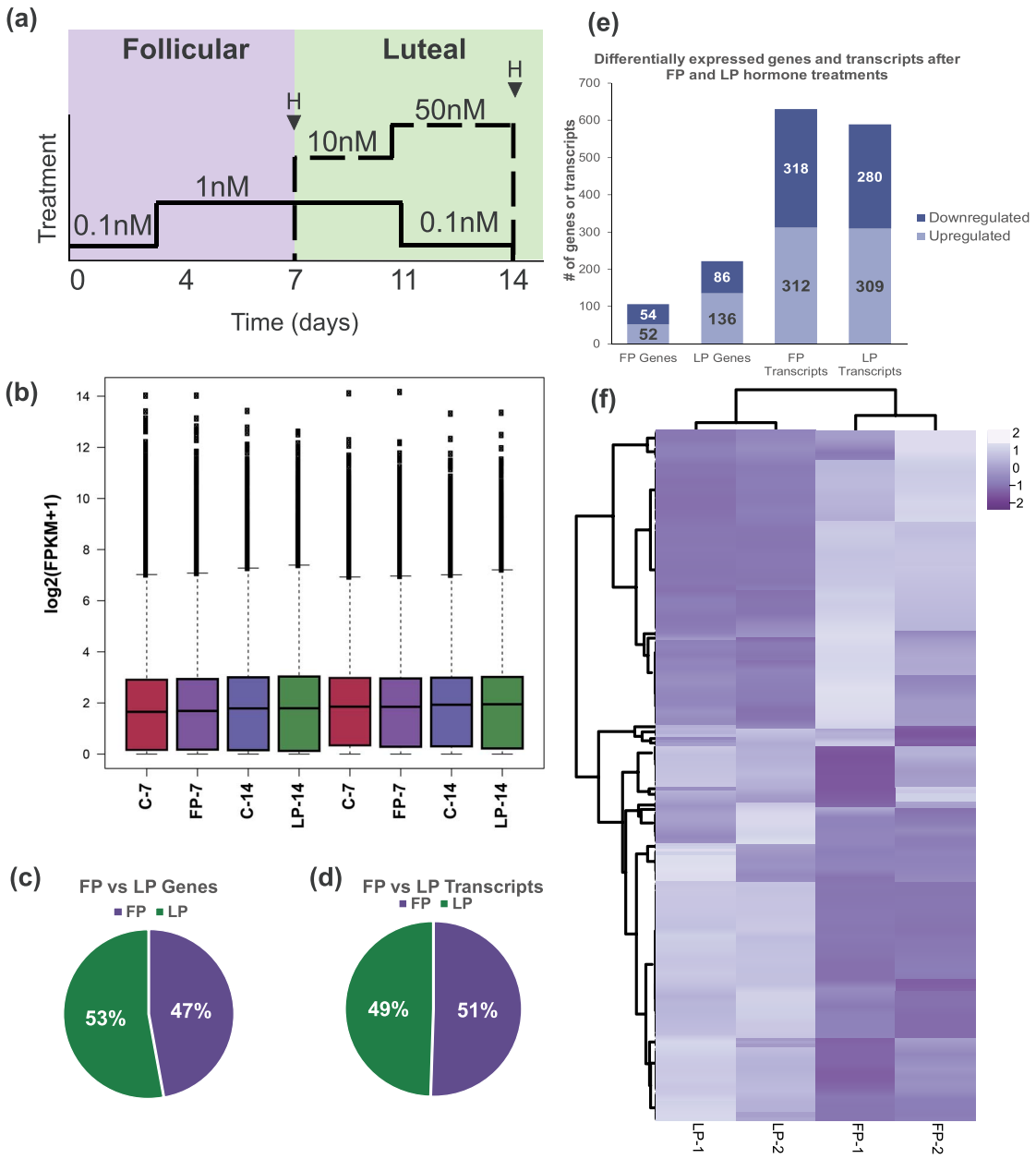


**Figure 1.** Distinct transcriptional profiles after follicular and luteal hormone treatments. Patient-matched engineered cervical tissue was exposed to FP hormones over 7 days (FP-7), follicular hormones for 7 days followed by LP hormones for 7 days (LP-14), or no hormones for 7 and 14 days (C-7, C-14) (A). RNA sequence analysis was performed on eight engineered cervical tissue samples exposed to follicular, luteal, or no hormone treatments. All samples showed consistent distribution of transcript fragments per kilobase per million (FPKM) counts, and the low median values demonstrate few very highly expressed genes (B). Of the 113 significant differentially expressed genes ($P < 0.05$, FC > 1.5), 47% were associated with FP treatments, and 53% with LP treatments, and at the transcript level, 197 (49%) were associated with the FP, while 193 (51%) were associated with the LP (C, D). Comparing FP and LP ECT to no hormone treatments showed abundant up- and down-regulated genes and transcript isoforms in response to FP (E2) or LP (E2 + P4) (E). An unsupervised cluster analysis of differentially expressed genes between samples grouped genes and samples into clusters based on expression similarity and revealed two distinct, reciprocal gene expression profiles for follicular and LP hormone treatments (F).

was analyzed as the fold change (FC) of FPKM value over the corresponding control FPKM value. ClustVis software [10] was used to perform unsupervised hierarchical cluster analysis on the resulting data, which grouped genes and samples by expression similarity.

## Gene ontology and functional annotation

Genes that were differentially expressed ($P < 0.05$, FC > 1.5) between FP and LP treatments, and between each treatment and corresponding control (C-7, C-14), were used to compare ontology terms using WebGestalt software [11] and DAVID functional annotation software. Webgestalt software assigned genes into overall summary categories for cellular components, biological processes, and molecular functions based on the number of genes associated with terms in each category. Resulting graphs are displayed as FC between FP and LP treatments based on proportion of genes assigned to each category. To further investigate cycle-phase specific ontology,

we performed a functional annotation analysis using the online DAVID bioinformatics database [12, 13] to compare gene ontologies, pathways, and key terms and group results of each into functional clusters. Genes most highly expressed in FP and LP treatments, as well as those up- or down-regulated by either FP or LP treatments relative to no hormone treatment control samples were compared. Gene ontology and pathway associations were considered significant when $P < 0.05$.

### Quantitative RT-PCR

Quantitative real-time PCR was performed on previously isolated RNA from engineered tissue samples from three patients, in four treatment conditions each (C-7, FP-7, C-14, and LP-14). Two of the patients were used in RNAseq analysis. Reactions for each condition were run in triplicate using StepOnePlus real-time PCR system (Thermo Fisher) and IDT gene expression assays (Integrated DNA Technologies, IL) to determine relative expression of each gene. Cycle thresholds (Cts) were calculated and normalized with ribosomal 18S (18s rRNA) gene expression as a control. Ct was placed at approximate level where increases in amplification were parallel between samples. Gene expression in engineered tissue was calculated using the $-\log2(\Delta\Delta Ct)$ method, relative to corresponding control for each treatment. All reactions were run 40 cycles (95 °C for 15 s, 60 °C for 1 min) after initial 3 min incubation at 95 °C.

## Results

### Data inspection and differential expression analysis

To determine transcriptional changes induced by cycling E2 and P4, RNAseq was performed on eight engineered cervical tissue (ECT) samples. Patient-matched conditions included control ECT—no hormone treatment, day 7 (C-7) ($n = 2$), day 7 of FP (FP-7) hormone treatments ($n = 2$), control day 14 (C-14) ($n = 2$), and day 14 of LP (LP-14) hormone treatments ($n = 2$) (Figure 1A). FastQC was used for quality check of reads and revealed high quality scores across all bases, with approximately 27 M human-specific mapping reads per sample. Accurate transcript assembly is dependent on quality of the read alignments, so to determine alignment quality achieved with TopHat2, StringTie (ST) [8] was used to merge the alignments of all samples and compare to reference genome hg_19 (UCSD). The "gffcompare" function was used to calculate base, exon, intron, intron chain, transcript, and locus-level sensitivity and precision, as well as total matching, missed, and novel counts. We found that samples aligned to the reference genome with 99.3–100% sensitivity and 95.2–98.9% precision, indicating high-quality alignment data for downstream analysis. Additionally, 3756/622 240 exons identified were novel, and 2843/63 216 loci were novel (Supplementary Table S1).

Transcript and gene abundances were measured in fragments per kilobase per million (FPKM). Large differences in distribution of FPKM values between samples can indicate problematic transcript alignment or assembly. To inspect FPKM distribution, abundance values were plotted as $\log2(FPKM + 1)$ for each. The resulting plot showed that all samples had a similar distribution (Figure 1B). Additionally, since most measurements fall below 5, these data suggest that only a small fraction of transcripts are expressed at very high levels in these samples.

We calculated fold change (FC) in gene and transcript FPKM values between treatments, including C-7 and FP-7; C-14 and LP-14; and between FP-7 and P-14 after normalization with the corresponding control for each. We found that 1775 genes were differentially expressed with $P < 0.05$ between FP-7 and LP-14 treatments, of which 499 were novel. Significance was defined as $P < 0.05$ and FC > 1.5. Of the known genes with $P < 0.05$, 123 genes had a fold change (FC) greater than 1.5 between FP-7 and LP-14 hormone treatments, with 53% of these genes expressed more highly in the LP group and 47% in the FP (Figure 1C). When comparing FP-7 samples to C-7, we found 116 significant differentially expressed genes, of which 51% were downregulated and 49% were upregulated (Figure 1D). LP-14 hormone treatments compared to C-14 revealed 222 differentially expressed genes with $P < 0.05$ and FC > 1.5. Of those, 39% were downregulated and 61% upregulated (Figure 1E). At the transcript level, we found 531 isoforms with significant differences in expression between FP-7 and LP-14 treatments, of which 139 were novel. Of the 390 known transcript isoforms, 51% were associated with FP-7 and 49% with LP-14 treatments (Figure 1D). Comparing FP-7 and LP-14 treatments with controls C-7 and C-14, we found 312 significantly upregulated and 318 significantly downregulated transcripts in response to FP-7 treatments, while LP-14 treatments resulted in 310 significantly upregulated and 281 significantly downregulated transcripts (Figure 1E). Tables of all significant differentially expressed genes and transcripts are available as Supplementary Data.

### Hierarchical clustering reveals two distinct gene profiles in ECT after follicular and LP hormone treatments

To visualize gene expression data, ClustVis software [10] was used to perform an unsupervised cluster analysis of significant differentially expressed genes between FP and LP treatments. Samples and genes were clustered based only on expression similarity, which resulted in two distinct gene expression profiles for FP and LP hormone treatments. The results are shown as a heatmap, with decreased expression represented by negative numbers in light purple and increased expression represented by positive numbers in dark purple (Figure 1F) The heatmap shows striking differences in gene expression in FP and LP treatments, with genes highly expressed in FP treatments downregulated in LP treatments and vice versa. This demonstrates that there are two distinctly different genomic profiles in engineered ectocervix based on menstrual cycle phase and highlights the importance of considering physiologic endocrine signaling in the study of reproductive tissues. Additionally, this serves as proof-of-concept that patient-matched engineered tissue, cultured with precisely controlled hormone concentrations can enable the discovery of significant biological insights from small sample sizes.

### Gene ontology analysis reveals differentially regulated cellular components, biological processes, and molecular functions after follicular and LP hormone treatments

To probe the phenotypes associated with differential gene expression after FP and LP treatments, we compared gene ontology (GO) terms associated with each gene set using Webgestalt online software [11], which assigned all significant genes into summary categories associated with cellular components (CCs), biological processes (BPs), and molecular functions (MFs). The proportions of genes assigned to each category after FP and LP treatments were compared, and results are reported as FC, highlighting the distinct differences in

**Table 1.** Functional gene clusters highly represented in FP samples.

| | Functional clusters | Genes |
|---|---|---|
| 1 | Lysosome, endosome, late endosome | LAMP2, SQSTM1, RAB17, RUFY1, AP5M1, SLC35F6 |
| 2 | Endosome, protein transport, transport | LAMP2, SLC44A2, RAB17, RUFY1, PARP11, CNIH1, AP5M1, SLC35F6, FXYD5, TIMM44 |
| 3 | Zinc finger, zinc iron binding, metal binding | ZC3H13, KDM2A, SQSTM1, SYTL4, RUFY1, LTA4H, HERC2, BAZ2B, FBXO11, NR2C1 |
| 4 | GTPase activity, Nucleotide binding, GTP binding | GTPBP2, PSMC6, EIF5, RAB17, TUBA1A, TIMM44 |
| 5 | Identical protein binding, chromatin binding, methylation | NONO, PSMC6, SQSTM1, RELA, HCFC1, STAT1, KLHDC3, WASF1, WDR74 |
| 6 | Transcription regulation, host-virus interaction, and isopeptide bond | MDFIC, RELA, HCFC1, STAT1, NONO, ZC3H13, RPS20, BAZ2B, TUBA1A, RBM25, UIMC1, NR2C1, CCNK, KDM2A |
| 7 | Cell adhesion, adherens junctions, cadherin binding involved in cell-cell adhesion | EIF5, HCFC1, STAT1 |
| 8 | Transit peptide, mitochondria | IMMT, TBRG4, HCFC1, SLC35F6, TIMM44 |
| 9 | Membrane, glycoprotein, transmembrane helix | ORAI2, ZC3H13, SLC44A2, TMEM214, IMMT, RUFY1, FXYD5, TIMM44, LHFPL2, LAMP2, CFL1, RAB17, ULBP2, SYTL4, RAET1L, HBEGF, AP5M1, CNIH1, SLC35F6 |
| 10 | Extracellular space, signal peptide, disulfide bond | LAMP2, SLC44A2, TMEM214, ULBP2, RAET1L, HCFC1, HBEGF, SLC35F6, FXYD5 |

representation of genes associated with CCs, BPs, and MFs based on the presence of FP or LP hormones. Cellular components more highly represented in FP samples included endosome, chromosome, vesicle, vacuole, and extracellular matrix, while ribosomes, mitochondria, cell projections, and Golgi apparatus were more highly represented in LP ECT (Figure 2A).

Biological processes highly associated with FP samples included antioxidant and electron carrier activity, nucleic acid binding, ion binding, and enzyme regulator activity, while those most associated with LP included molecular adaptor, structural molecule, and transferase activity (Figure 2B). Molecular functions associated with FP samples included growth, proliferation, multi-organism process, reproduction, cellular component biogenesis, developmental process, biological regulation, cell localization and communication, and multicellular processes, while MFs associated with LP treatments included response to stimulus (Figure 2C). These results indicate distinct differences in active CCs, BPs, and MFs in ECT dependent on menstrual cycle phase hormones.

## Comparison of genes and functional clusters associated with follicular and LP hormone treatments

To further investigate the profiles summarized by Webgestalt, a DAVID functional annotation analysis was performed, which grouped GO terms from CCs, BPs, and MFs, in addition to KEGG pathways, key processes, and protein-protein interactions into functional clusters most over-represented in each gene set. This resulted in significantly different functional clusters between the FP and LP treatment groups. The FP samples were characterized by genes associated with endosomes and protein transport. FP genes also included many associated with zinc and metal binding. Zinc-binding is involved in transcription and cell cycle, so unsurprisingly many genes associated with chromatin binding and transcription were also highly expressed in FP samples. Genes that play a role in cell adhesion were also enriched, as well as a number of glycoproteins and proteins localized to the membrane or extracellular space (Table 1).

Genes highly expressed in the LP were associated with Golgi apparatus and mitochondria. The Trans-Golgi Network sorts membrane and secreted proteins, many of which were also highly expressed after LP treatments (Table 2). Additionally, the LP samples showed high expression of genes associated with hydrolase and protease activity, as well as chemokine signaling and B and T-cell receptor signaling. These data represent two distinctly different functional profiles in engineered tissue dependent on FP or LP hormones. However, it cannot be assumed that genes associated with the FP are E2-regulated, as it is possible that expression in LP samples was inhibited by P4 signaling, rather than induced by E2 signaling in FP samples, and vice versa.

To further investigate FP and LP gene and functional associations, categories of genes that were up- or down-regulated by each treatment were compared relative to samples that were not exposed to hormones. The FP treatments (0.1–1 nM E2) upregulated genes involved in zinc and DNA-binding, as well as membrane components and protein transport. In contrast, the E2 treatments downregulated lipid metabolism and degradation genes. Genes that repress transcription were also downregulated, as were several secreted glycoproteins (Table 3).

LP treatments (1–0.01 nM E2 and 10–50 nM P4) resulted in increased expression of genes involved in RNA binding, splicing, and processing. Genes associated with cytokine and MAPK signaling were also upregulated by LP hormones. Additionally, actin-binding genes involved in cytokinesis and cell motility were upregulated. In contrast, LP treatments downregulated genes associated with zinc-binding and transcription, and genes involved in serine/threonine kinase activity and ATP-binding. Finally, glycosylated and secreted proteins such as *OVGP1* and others were downregulated (Table 4). These data suggest that the dynamic functional profiles for ECT across the menstrual cycle are a result of both complementary and contrasting hormone-receptor mediated events induced by E2 and P4.

## Validation of RNA sequence data by qRT-PCR

To validate RNA sequence data, qRT-PCR was performed probing expression of selected genes from differential expression analysis or genes of interest related to functional categories in ECT after FP-7 ($n = 3$) or LP-14 ($n = 3$) treatments compared to no hormone

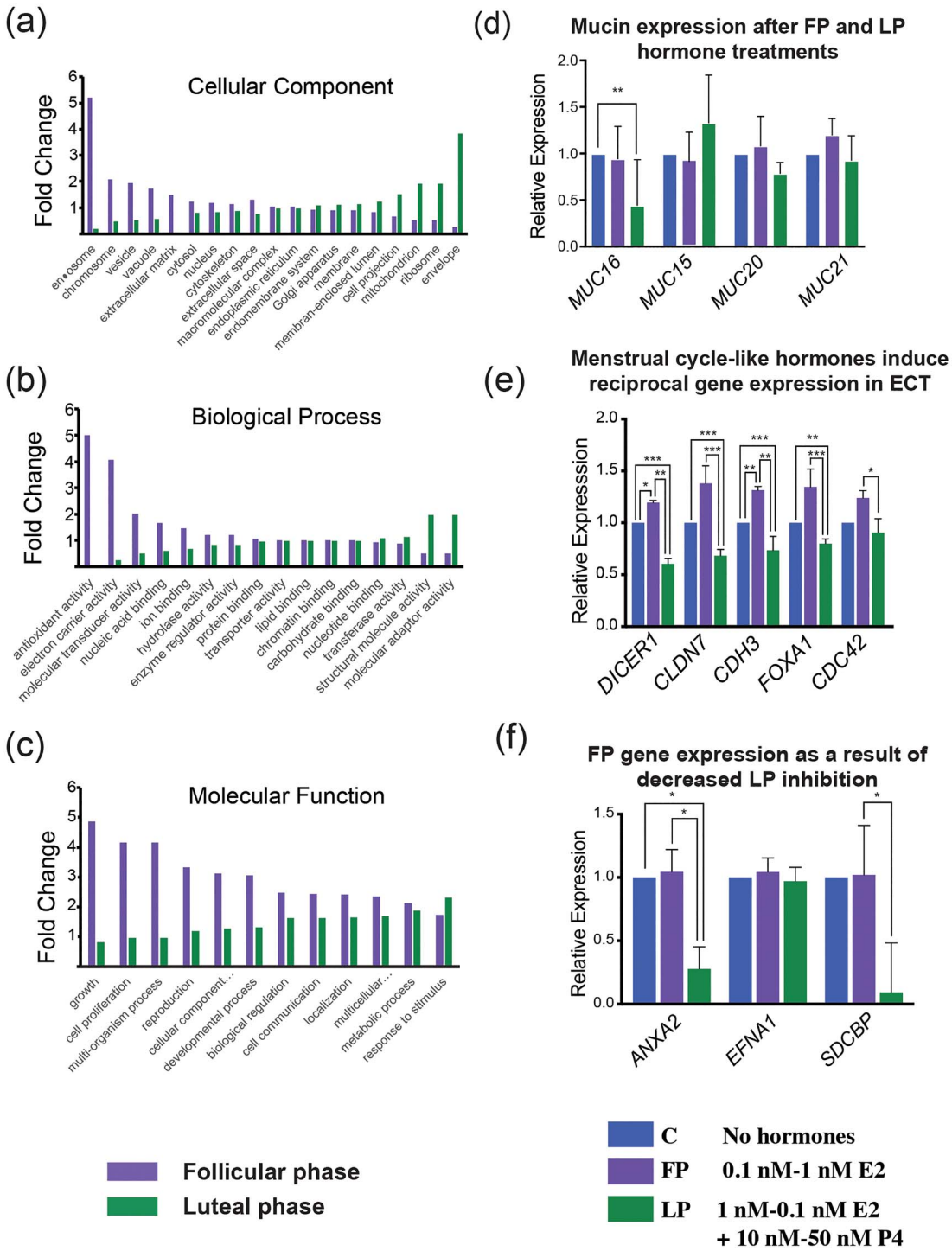


**Figure 2.** Gene ontology analysis and qPCR validation of hormone responsive genes. Proportions of genes assigned to ontology categories of cellular components (A), biological processes (B), and molecular functions (C) revealed significantly different profiles based on FP or LP hormones. Gene expression of selected genes in ECT was analyzed by qRT-PCR ($n = 3$ per group). Resulting data are shown as expression relative to samples that received no hormone treatments. DICER1, CLDN7, CDH3, FOXA1, and CDC42 were upregulated in the FP samples; in contrast, all were significantly downregulated in the LP samples (E). Expression of ANXA2 and SDCBP did not seem to be estrogen-responsive but were significantly downregulated in the LP samples. EFNA1 on the other hand was not significantly different between the groups and the control (F). Expression of MUC21, MUC20, MUC15, and MUC16 was identified in RNAseq data. We confirmed by PCR and found that MUC 16 was significantly downregulated in LP samples. MUC20 and 21 were also downregulated in LP samples, while MUC15 was upregulated, but this did not reach statistical significance (D). ($^*P < 0.05$, $^{**}P < 0.001$, $^{***}P < 0.0001$).

treatments C-7 ($n = 3$) and C-14 ($n = 3$). Previously, we had investigated *CLDN1* and *CDH1* expression in ECT, and found no significant difference in expression across the menstrual cycle (data not shown); however, since genes expressed more highly in FP samples were associated with cell adhesion, adherens junctions, and cell communication, we analysed read-data of genes previously shown to

**Table 2.** Functional gene clusters highly represented in LP samples.

| | Functional clusters | Genes |
|---|---|---|
| 1 | B-cell receptor signaling pathway, T-cell receptor signaling pathway, chemokine signaling pathway | SOS1, IKBKB, VAV1 |
| 2 | Golgi apparatus, Golgi membrane | TGOLN2, NDFIP1, ATP8B1, SPPL2B, TRIP10, ITM2B, YKT6 |
| 3 | Transit peptide, mitochondria | USP30, ECH1, MRPL49, ACAD8, C6ORF203, YKT6, MDH2, GLRX, DDX6, MTFR1 |
| 4 | Hydrolase, protease, thiol protease | USP30, USPL1, SENP1, ATP8B1, SPPL2B, RNGTT, DDX6 |
| 5 | Cell projection, cytoskeleton | CSPP1, ITGB1BP1, ABI2, CEP164, TRIP10, CENPJ, FMR1, ATP8B1, |
| 6 | Intracellular, transferase, proton acceptor, kinase, ATP-binding, protein phosphorylation | IRAK4, MAP4K5, ZNF738, SOS1, PIK3C3, EDA2R, ZNF816, IKBKB, VAV1, GYS1, YKT6, RNGTT, GLRX, FDFT1, ACAD8, MDH2, DDX6 |
| 7 | Protein transport | SFT2D1, FMR1, ATP8B1, YKT6, NUP43, GLRX, ITGB1BP1 |
| 8 | Nucleolus, poly(A) RNA binding, Ubl conjugation | USP30, SP100, FMR1, NDFIP1, IKBKB, UTP14A, DDX6 |
| 9 | Membrane, transmembrane helix | TGOLN2, SFT2D1, USP30, TMEM187, NDFIP1, ATP8B1, SPPL2B, EDA2R, ZNF816, ITM2B, PAAF1, FDFT1 |
| 10 | Secreted, extracellular region, glycoprotein | TGOLN2, EFEMP2, SPPL2B, EDA2R, PCYOX1, IGFBP2, ITM2B, MDH2, MIA, NDFIP1 |

**Table 3.** Functional gene clusters upregulated and downregulated by FP E2.

| | Functional clusters | Upregulated genes |
|---|---|---|
| 1 | Metal-binding, zinc, zinc-finger | THAP7, POLI, POGZ, CNDP2, EPS15L1, ZNF669, ATP11C, NEK11, MAN2A2, ZNF706, RNF141, SLC25A25, ZNF146, ZSCAN32, NARFL, ZZZ3, GNAS, PGGT1B, PHF6, RNF13 |
| 2 | Nucleus, DNA-binding, transcription regulation | THAP7, POLI, POGZ, TAF8, EPS15L1, ZNF669, COMMD9, NEK11, MSL3, ZNF706, STAT4, H2AFV, MED17, ZNF146, ZSCAN32, NELFA, RPP30, ZZZ3, PHF6, RNF13 |
| 3 | Cell membrane, plasma membrane | SLC25A30, MAN2A2, RGMB, RNF141, SLC25A25, EPS15L1, GNAS, ATP11C, ADD3, SLC39A1, RNF13, RNF141, MED17, EPS15L1, HSPE1, GNAS, ADD3, SLC39A1 |
| 4 | Transport, endoplasmic reticulum, transmembrane region | SLC25A30, MAN2A2, SLC25A25, SLC39A1, ATP11C, RNF13, RABEP1, COMMD9 |
| | Functional clusters | Downregulated genes |
| 1 | Lipid degradation, lipid metabolism, hydrolase | DDHD2, LIPG, MUS81, PLA2G4D |
| 2 | Transcription, nucleus | CDK1, ANKRD54, KHDRBS3, GEMIN2, HOXA4, VHL, MUS81, FSTL3, HMG20B, GTF3C4, SUDS3, ATF7IP2, MLF1 |
| 3 | Lipoprotein, cell membrane, plasma membrane | RAB3D, PDPN, IFITM2, LIN7B, ARAP3, ALDH3B1 |
| 4 | Membrane, transmembrane helix | RAB3D, VHL, PDPN, IFITM2, MTX1, C17ORF62, LIN7B, TMEM42, ICMT, ALDH3B1, PRRG1, LRCH3, ROBO3, ARAP3, MANBAL, PLA2G4D |

be expressed in the ectocervix involved in tight junction and adherens junctions, claudin 7 (*CLDN7*) and cadherin 3 (*CDH3*), that may interact with those genes identified in our analysis. Claudin7 and *CDH3* (tight junctions, cell adhesion), *FOXA1* (transcription), and *DICER1* (RNA binding) were all significantly upregulated in FP and downregulated in LP ECT (Figure 2E).

## Mucin expression in ECT after FP and LP hormone treatments

Mucins are heavily glycosylated proteins that are an important part of the chemical and physical barrier in ectocervical tissue during homeostasis and can also play a role in cancer progression and metastasis [14–16]. Previously, *MUC1, MUC4*, and *MUCL1* gene expression have been reported in ectocervical tissue [17, 18], and we confirmed that engineered tissue mimicked expression and localization patterns of mucins observed in PT (BOR_1). Read data from RNA sequence analysis identified mRNA transcripts for additional mucins not previously reported in the cervix. We found that *MUC15, MUC16, and MUC3A* were abundantly expressed in our samples, while *MUC12, MUC13, MUC19, MUC20, MUC22, MUC5B,* and *MUC6* were expressed in low amounts in ectocervical epithelial cells. Additionally, these data suggest that many mucin genes may be regulated by ovarian hormones, as we observed higher FPKM values of *MUC16*, and *MUC3A* in FP samples and higher FPKM values of *MUC15* in LP samples. We validated expression of four mucin genes by qRT-PCR and found that *MUC16* was significantly downregulated in the LP (Figure 2D). To our knowledge, this is the

**Table 4.** Functional gene clusters upregulated and downregulated by luteal hormones E2 and P4.

| | Functional clusters | Upregulated genes |
|---|---|---|
| 1 | K Homology domain, RNA binding | FUBP3, SNRPA, QKI, CPSF4, AKAP1, JUNB |
| 2 | Kinase, transferase | SEPHS2, BCR, UBE2Z, DHDDS, AK1, EPHB2, CDKN1C, GALK2, POMT1, ASMTL, FUT2, AKAP1, POFUT1, LIPT1, MAP3K11 |
| 3 | protein kinase binding, apoptosis, herpes simplex infection, hepatitis c | PEA15, UBE2Z, TICAM1, JTB, TP53, IGFBP3, TRAF3, SRSF8 |
| 4 | Transcription, nucleus, DNA-binding | SNAPC5, SUPT3H, SSBP3, FUBP3, FOXK2, TP53, HMGA1, JUNB, CBFB, DVL1, ESRRA, MED25, RHOQ, RGMA, VEGFA, AGRN, SMARCA4, ZBTB47, ZNF282, ZNF646, ZNF35, PHC2 |
| 5 | Oxidoreductase, NAD(P)-binding domain | GPD1L, RRM2, PHGDH, MICAL1, SESN2, BDH1, GFOD2 |
| 6 | MAPK signaling pathway, neuron apoptotic process | IL1R1, TP53, TGFB2, MAP3K11 |
| 7 | Poly(A) RNA binding, mRNA splicing, RNA processing | FUBP3, EFTUD2, SRSF8, FSCN1, SNRPA, QKI, CPSF4, UNK, MYH9, AKAP1, ZNF346, CDC42EP4, KAP1, JUNB |
| 8 | Lysosome, lysosomal membrane, endosome | TMEM175, LAMTOR1, ABCA2, KCNK1, TNFAIP1, TRAF3, FUCA1, MFSD12, AP1B1 |
| 9 | Actin-binding, cadherin-binding | IMPACT, FSCN1, MICAL1, MYH9, EMD |
| 10 | Apoptosis | PEA15, UBE2Z, TICAM1, JTB, TP53, IGFBP3, TRAF3, TNFAIP1, ANAPC11, STUB |
| 11 | Pathways in cancer, positive regulation of gene expression | BCR, VEGFA, TP53, TGFB2, TRAF3, DVL1, FUBP3, QKI |
| | Terms | Downregulated genes |
| 1 | Zinc, zinc-finger, transcription, metal binding, nucleus | GALNT3, KLF6, CNOT8, GEN1, PPM1A, ZNF25, RNF217, CLYBL, ZNF615, TRIM38, ZNF671, DNAJC24, ZSCAN32, ZIK1, NEK4 |
| 2 | Mitochondrial translation termination, elongation | MRPL2, CKMT1B, PTCD3, CLYBL, DAP3, CBR4, SLC44A1 |
| 3 | WD40 repeat | WRAP73, WDR77, WDR20 |
| 4 | Mitosis, cell division, cell cycle | HAUS3, BUB1B, NEK4 |
| 5 | cell junction, synapse | GRIP1, PSD3, LGI3, GID8 |
| 6 | serine/threonine kinase, ATP-binding | ALPK1, CKMT1B, OPLAH, RFC2, BUB1B, NEK4 |
| 7 | Membrane, transmembrane | GALNT3, SLC44A1, CKMT1B, GRIP1, FUT10, PSD3, PPM1A, TIRAP, RNF217, CLYBL, TMEM67, PRRG1, DNAJC24, HECTD4, ERC1, PLA2R1, CMTM4, MEST |
| 8 | Glycosylation, secreted, glycoprotein | GALNT3, TMEM67, OVGP1, FUT10, LGI3, TCTN1, PLA2R1, MEST, LGI3 |

first time mRNA transcripts for these 10 mucins have been identified in ectocervical epithelium. Differential expression suggests that dynamic expression of mucins across the menstrual cycle may result in specific mucosal functional characteristics, which has implications for both infection and reproduction.

## Discussion

To determine changes occurring in ectocervical epithelium across the course of the menstrual cycle, we analyzed RNA sequence data of engineered ectocervical tissue (ECT) after follicular and luteal phase (FP, LP) hormone treatments, and compared to untreated ECT from the same patients. Genes and transcripts most highly associated with follicular and LP treatments were identified, enabling the identification of distinct transcriptional profiles for each phase of the cycle. Additionally, we further probed the role of each of the phase-specific hormone combinations (0.1–1 nM E2, or 1–0.1 nM E2 + 10–50 nM P4) to investigate whether genes and functional clusters associated with each phase were expressed in response to the dominant steroid hormone associated with each phase (FP-E2, LP-P4) or simply a result of decreased inhibitory signaling from the opposing receptor. Our studies indicate that hormone receptors in ectocervical epithelium act in complimentary and antagonistic ways to precisely regulate molecular and biological functions in response to cycling levels of E2 and P4. Indeed, many of the same genes and processes upregulated in the FP are downregulated in the LP and vice versa, highlighting the complementary signaling that occurs between the receptors to precisely regulate gene expression based on menstrual cycle phase.

As is the case in other female reproductive tissues, we found that FP ECT samples were associated with genes involved in proliferation, transcription, chromatin-binding, and cell cycle regulation. Additionally, genes associated with membrane components, cell adhesion, and secreted and transmembrane glycoproteins were also more highly expressed in FP samples. Interestingly, when we compared FP samples to control samples, we found that E2 seems to both upregulate and downregulate the same categories of genes. For example, the four most upregulated functional clusters in response to E2 were (1) metal binding, zinc, zinc finger; (2) nucleus, DNA-binding, transcription; (3) membrane, cell membrane, plasma membrane; and (4) transport and transmembrane region. The most highly downregulated cluster by E2 was lipid degradation and metabolism. Other downregulated clusters included transcription and nucleus, membrane and transmembrane, as well as extracellular and secreted glycoproteins. However, upon further analysis of the genes associated with each category, we found that those which

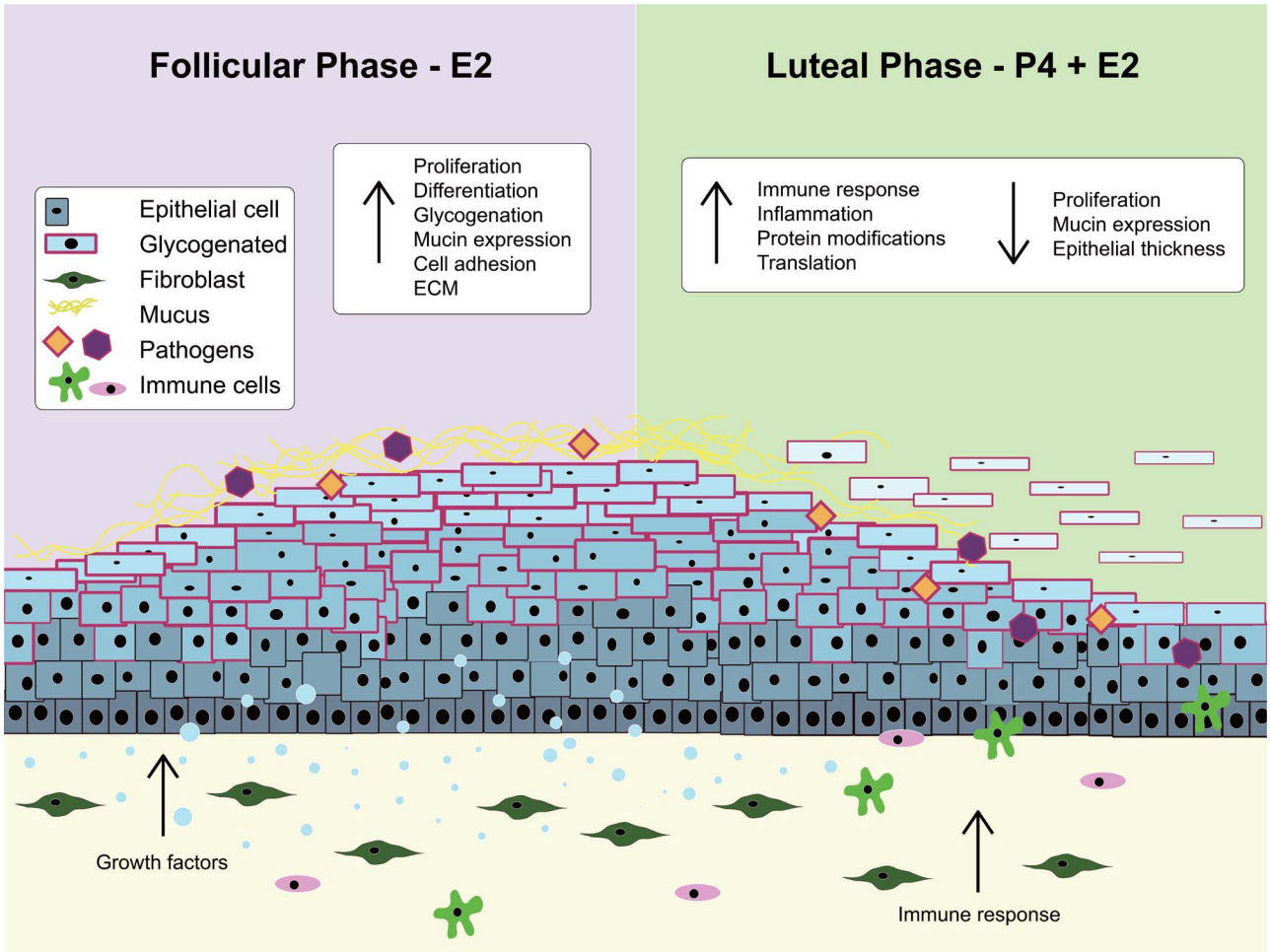


**Figure 3.** Hormonal regulation of ectocervical mucosa across the menstrual cycle. A number of barrier properties are upregulated during the FP of the menstrual cycle, including proliferation, differentiation, glycogenation, mucin expression, adhesion, and growth factor signaling. LP treatments were characterized by increased immune response, inflammation, protein modifications, and decreased proliferation, mucin expression and epithelial thickness, indicating that ectocervical tissue may be more vulnerable to infection in the LP of the menstrual cycle.

were upregulated had distinctly different functions from the genes in the same categories which were downregulated. For example, E2 downregulates Activating Transcription Factor 7 Interacting Protein 2 (*ATF7IP2*), which functions to repress transcription; thus, E2 indirectly upregulates transcription in addition to the direct upregulation of many genes directly involved in transcription and DNA-binding.

In contrast, LP samples were more highly associated with RNA-binding and signaling. When comparing LP to FP samples, we found LP samples more highly expressed genes involved in insulin-like growth factor signaling, chemokine signaling, as well as B and T cell signaling. When comparing LP samples to control samples, a number of additional signaling pathways were found to be upregulated, such as MAPK, TGFB2, NFKB, and TNF$\alpha$. LP gene expression also was highly associated with immune response and inflammation, as well as a number of genes involved in viral infection. For example, P4 upregulated expression of *FOXK2*, which can bind to HIV motifs and regulate viral transcription. The most enriched transcript in luteal vs. FP samples was *TLR3*, a key component to host immune response. Additionally, *IRF7* was found to be highly expressed in LP samples and has been previously shown to play a role in activating transcription of virus-inducible genes. Additionally, we report the expression of an additional 10 mucins not previously identified in ectocervical tissue, many of which appear to be hormonally regulated. Mucins play a vital role in both the barrier and reproductive properties of cervical tissue and are increasingly becoming a target for drug delivery to mucosal tissues.

Overall, our data support the theory that there is a "window of vulnerability" in the ectocervix for infection based on the menstrual cycle phase, with the high progesterone LP being the most vulnerable. This is the first study to characterize differential gene expression throughout the menstrual cycle by RNAseq analysis in ectocervical tissue (summarized in Figure 3) and has identified a number of new hormonally regulated genes in this tissue that may influence susceptibility to infection or could potentially be used as targets for drug delivery or preventive medicine.

## Supplementary data

Supplementary data are available at *BIOLRE* online.

## Acknowledgements

The authors thank the University of Chicago Genomics Facility—Knapp Center for Biomedical Discovery for next generation sequencing.

## Conflict of interest

The authors have declared that no conflict of interest exists.

## References


1. Wira CR, Ghosh M, Smith JM, Shen L, Connor RI, Sundstrom P, Frechette GM, Hill ET, Fahey JV. Epithelial cell secretions from the human female reproductive tract inhibit sexually transmitted pathogens and Candida albicans but not lactobacillus. *Mucosal Immunol* 2011; **4**: 335–342.
2. Hickey DK, Patel MV, Fahey JV, Wira CR. Innate and adaptive immunity at mucosal surfaces of the female reproductive tract: Stratification and integration of immune protection against the transmission of sexually transmitted infections. *J Reprod Immunol* 2011; **88**: 185–194.
3. Thurman AR, Chandra N, Yousefieh N, Zalenskaya I, Kimble T, Asin S, Rollenhagen C, Anderson SM, Herold B, Mesquita PM, Richardson-Harman N, Cunningham T et al. Comparison of follicular and luteal phase mucosal markers of HIV susceptibility in healthy women. *AIDS Res Human Retrovir* 2016; **32**:547–560.
4. Irvin SC, Herold BC. Molecular mechanisms linking high dose medroxyprogesterone with HIV-1 risk. *PLoS One* 2015; **10**:e0121135.
5. Deese J, Masson L, Miller W, Cohen M, Morrison C, Wang M, Ahmed K, Agot K, Crucitti T, Abdellati S, Van Damme L. Injectable progestin-only contraception is associated with increased levels of pro-inflammatory cytokines in the female genital tract. *Am J Reprod Immunol* 2015; **74**:357–367.
6. McKinnon K, Woodruff T, Getsios S, Sensharma R, Ravix J, Williams C. Development of Human Ectocervical Tissue Models with Physiologic Endocrine and Paracrine Signaling. *Biol Reprod* 2020; (in press). DOI: 10.1093/biolre/ioaa068.
7. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**:R36.
8. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015; **33**:290–295.
9. Frazee AC, Pertea G, Jaffe AE, Langmead B, Salzberg SL, Leek JT. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol* 2015; **33**:243–246.
10. Metsalu T, Vilo J. ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res* 2015; **43**:W566–W570.
11. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res* 2017; **45**, W130–W137. https://doi.org/10.1093/nar/gkx356.
12. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**:44–57.
13. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009; **37**:1–13.
14. Amini A, Masoumi-Moghaddam S, Morris DL. Mucins and Tumor Biology. In: Amini A, Masoumi-Moghaddam S, Morris DL (eds.), *Utility of Bromelain and N-Acetylcysteine in Treatment of Peritoneal Dissemination of Gastrointestinal Mucin-Producing Malignancies*. Cham: Springer International Publishing; 2016: 43–61.
15. Rao CV, Janakiram NB, Madka V, Kumar G, Scott EJ, Pathuri G, Bryant T, Kutche H, Zhang Y, Biddick L, Gali H, Zhao YD et al. Small-molecule inhibition of GCNT3 disrupts mucin biosynthesis and malignant cellular behaviors in pancreatic cancer. *Cancer Res* 2016; **76**: 1965–1974.
16. Liu ZZ, Xie XD, Qu SX, Zheng ZD, Wang YK. Small breast epithelial mucin (SBEM) has the potential to be a marker for predicting hematogenous micrometastasis and response to neoadjuvant chemotherapy in breast cancer. *Clin Exp Metastas* 2010; **27**:251–259.
17. Gipson IK, Ho SB, Spurr-Michaud SJ, Tisdale AS, Zhan Q, Torlakovic E, Pudney J, Anderson DJ, Toribara NW, Hill JA 3rd. Mucin genes expressed by human female reproductive tract epithelia. *Biol Reprod* 1997; **56**:999–1011.
18. Ayehunie S, Islam A, Cannon C, Landry T, Pudney J, Klausner M, Anderson DJ. Characterization of a hormone-responsive organotypic human vaginal tissue model: morphologic and immunologic effects. *Reprod Sci* 2015; **22**:980–990.