Molecular Signatures of Immune Activation and Epithelial Barrier Remodeling Are Enhanced during the Luteal Phase of the Menstrual Cycle: Implications for HIV Susceptibility

Kenzie Birse, Kelly B. Arnold, Richard M. Novak, Stuart McCorrister, Souradet Shaw, Garrett R. Westmacott, Terry B. Ball, Douglas A. Lauffenburger, Adam Burgener

National Lab for HIV Immunology, JC Wilt Infectious Disease Research Centre, Public Health Agency of Canada, Winnipeg, Manitoba, Canada; Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada; Department of Biological Engineering and Center for Gynepathology Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; Mass Spectrometry Core Facility, National Microbiology Laboratory, Winnipeg, Manitoba, Canada; Department of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA; Department of Medical Microbiology, University of Nairobi, Department of Medical Microbiology, Nairobi, Kenya; Unit of Infectious Diseases, Department of Medicine Salna, Center for Molecular Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden; Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

ABSTRACT

The variable infectivity and transmissibility of HIV/SHIV has been recently associated with the menstrual cycle, with particular susceptibility observed during the luteal phase in nonhuman primate models and \textit{ex vivo} human explant cultures, but the mechanism is poorly understood. Here, we performed an unbiased, mass spectrometry-based proteomic analysis to better understand the mucosal immunological processes underpinning this observed susceptibility to HIV infection. Cervicovaginal lavage samples \((n = 19)\) were collected, characterized as follicular or luteal phase using days since last menstrual period, and analyzed by tandem mass spectrometry. Biological insights from these data were gained using a spectrum of computational methods, including hierarchical clustering, pathway analysis, gene set enrichment analysis, and partial least-squares discriminant analysis with LASSO feature selection. Of the 384 proteins identified, 43 were differentially abundant between phases \((P < 0.05, \geq 2\text{-fold change})\). Cell-cell adhesion proteins and proteases were reduced, and leukocyte recruitment \((\text{interleukin-8 pathway, } P = 1.41\text{E}{-5})\) and extravasation proteins \((P = 5.62\text{E}{-4})\) were elevated during the luteal phase. LASSO/PLSDA identified a minimal profile of 18 proteins that best distinguished the luteal phase. This profile included cytoskeletal elements and proteases known to be involved in cellular movement. Gene set enrichment analysis associated CD4\(^{+}\) T cell and neutrophil gene set signatures with the luteal phase \((P < 0.05)\). Taken together, our findings indicate a strong association between proteins involved in tissue remodeling and leukocyte infiltration with the luteal phase, which may represent potential hormone-associated mechanisms of increased susceptibility to HIV.

IMPORTANCE

Recent studies have discovered an enhanced susceptibility to HIV infection during the progesterone-dominant luteal phase of the menstrual cycle. However, the mechanism responsible for this enhanced susceptibility has not yet been determined. Understanding the source of this vulnerability will be important for designing efficacious HIV prevention technologies for women. Furthermore, these findings may also be extrapolated to better understand the impact of exogenous hormone application, such as the use of hormonal contraceptives, on HIV acquisition risk. Hormonal contraceptives are the most widely used contraceptive method in sub-Saharan Africa, the most HIV-burdened area of the world. For this reason, research conducted to better understand how hormones impact host immunity and susceptibility factors important for HIV infection is a global health priority.

The global HIV incidence rates for women remain high, so much so that every minute a young woman becomes infected with HIV \((1)\). With infection rates twice as high for women aged 15 to 24 compared to their age-matched male counterparts, it is clear that young women are more susceptible to HIV infection \((1, 2)\). Understanding biological contributors to HIV risk in women is therefore important for public health. Although still a topic of great debate there is a growing body of evidence that both endogenous and exogenous ovarian hormones, including estradiol and progesterone, are associated with increased HIV acquisition in women \((3–7)\). Studies of nonhuman primates have shown that SHIV infection occurs more frequently during the progesterone-dominant luteal phase of the menstrual cycle \((8, 9)\), and recent \textit{in vitro} explant cultures of human cervical tissue have also shown that HIV infects explants collected during the luteal phase more productively than those collected during the follicular phase \((10)\). Furthermore, progestins, which are synthetic progesterone derivatives commonly used in injectable hormonal contraceptives, have been shown to increase HIV acquisition risk by as much as
1.5- to 2.0-fold based on the latest meta-analysis (7, 11). Clearly, these results indicate a link between ovarian hormones and HIV risk, and understanding the mechanisms underpinning these observations may have important implications for sexual and reproductive health.

During heterosexual intercourse, HIV first encounters a mucosal surface and must overcome the defensive barriers of this surface to establish an infection. The stratified, squamous epithelial layer of the lower genital tract and its enveloping mucosal fluid represent the physical defense barriers meant to protect from invading microorganisms, and within the mucosal fluid are antimicrobial factors secreted by epithelial cells and various other immune cells which contribute to the molecular defense barrier. Estradiol and progesterone are ovarian hormones which vary over the course of the menstrual cycle, and it is their dynamically changing levels that govern various biological processes which are necessary for successful reproduction, including ovulation, implantation, and menstruation when conception does not occur. The changing levels of these endogenous hormones are also known to induce transformations in the defensive barriers of the female genital tract, including changes in mucus viscosity, epithelial barrier thickness (particularly within the stratum corneum), and disruptions in resident microflora typically dominated by Lactobacillus species (12). This may have particular relevance to HIV as the lower pH offered by Lactobacillus microflora can inhibit infection, and hormone-induced epithelial thinning and barrier disruption may increase the risk of HIV finding its target cells. Immune cell numbers are also known to vary over the menstrual cycle, including natural killer cells, neutrophils, macrophages, and lymphocyte aggregates, all of which are found at higher levels during the luteal phase in both the peripheral blood and within the upper female genital tract in the endometrium (13, 14). Immune cell numbers are believed to fluctuate much less in the lower female genital tract (15); however, changes in activation status/cellular phenotype and/or immune cell localization over the course of the menstrual cycle within this compartment, the area where HIV founder populations are thought penetrate the mucosal barrier (16), are not known.

Ovarian hormones may also have an impact on local inflammation and the expression of innate immune factors in mucosa. There are many innate immune factors present in mucosal fluid known to affect HIV infectivity in vitro, including defensins, elastase, elafin, serpins, mucosal IgA, and others (17–20), and several of these are associated with HIV acquisition risk (18, 21). For example, elevated antimicrobial factors such as alpha-defensins and SLPI have been associated with an increased risk of HIV acquisition in the foreskins of men (21), and increased levels of alpha-defensins and LL-37 in cervicovaginal secretions have been associated with increased HIV risk in women (22). In contrast, anti-inflammatory factors such as antiproteases have been associated with reduced acquisition among highly exposed sex workers (18, 23). Several of these antiviral factors have shown a relationship with the menstrual cycle and the corresponding ovarian hormones dominant during each phase, including an elevation of serpins during the follicular phase (24), and a reduction of SLPI and defensins during the ovulatory/early luteal phase (25, 26). Although there is a body of evidence supporting an association between progesterone and menstrual cycle phase with altered host mucosal innate immunity, a comprehensive analysis has never been performed and represents a major gap in knowledge. Therefore, a better understanding of the relationship between ovarian hormone levels and host immunity of the lower genital tract over the menstrual cycle may shed light on potential susceptibility mechanisms.

Standard analytical tools, which study single factors in isolation to assess host immunity, may undervalue complex biological events and interactions important for host immunity. Indeed, mucosal secretions are exceptionally complex, containing many hundreds of unique factors important for both innate immune defense, as well as inflammatory and anti-inflammatory processes (23, 27). It therefore follows that factors that modulate risk of HIV infection are likely complex, involving a network of multiple interacting components. High-throughput technologies, such as mass spectrometry-based proteomics, may help uncover novel biological processes not observable by traditional approaches. When coupled with multivariate data-driven modeling, this method has the capacity to provide a wide range of insight into biological systems, from diagnosis and classification of disease states to new insights into biological mechanisms. Partial least-squares deterministic analysis (PLSDA) is one such technique and can be used to determine patterns of features (in this case protein expression) that best distinguish between groups defined by a binary variable (in this case, the menstrual cycle phase). This approach has been successfully used to predict in vivo apoptotic and proliferative phenotypes in the mouse intestine after exposure to tumor necrosis factor alpha (28), as well biological events shown to occur during cytokine-associated mucosal inflammation (29–31).

We utilized here an unbiased proteomics approach to characterize the relationship between ovarian hormones and host mucosal immunity by examining mucosal secretions collected during different phases of the menstrual cycle. We report the distinguishing features of the luteal phase through hierarchical clustering and data-driven analysis techniques, which include enhanced molecular pathways involved with tissue remodeling and immune cell movement/migration, with a concomitant reduction of cell-cell adhesion and anti-inflammatory pathways. This is the first integrated systems biology study to examine the differences in the proteome of the human female genital tract during the hormonally controlled menstrual cycle and provides new information on mucosal processes which may be important for HIV infection.

MATERIALS AND METHODS

Study population and clinical information. The women who participated in this study provided written, informed consent. The study was approved by the human subjects committee of the University of Illinois at Chicago and Research Ethics Board of the University of Manitoba. All of the participants underwent testing for HIV, bacterial vaginosis, Trichomonas vaginalis, Neisseria gonorrhoeae, and Chlamydia trachomatis at the time of sample collection. Samples positive for any of these tests were excluded from the study. Participants were between the ages of 18 to 47, were not on any form of hormonal contraception, and abstained from sexual intercourse for at least 24 h prior to sample collection. Participants also abstained from using any kind of vaginal medication/creams or douching for at least 24 h prior to sample collection. All other relevant clinical information has been provided in Table S1 in the supplemental material.

CVL collection. For all participants, a speculum without the use of lubricant was inserted into the vagina, and the cervix was located. Four cotton tipped brushes were then used to swab the posterior, lateral, frontal, and cervical areas of the vaginal vault. These swabs were used for the standard testing of the sexually transmitted infections (STIs) described above. Cervicovaginal lavage (CVL) was then obtained by instilling 10 ml of saline solution over the surface of the vaginal vault and ectocervix. The
saline lavage was then redrawn (8 to 10 ml) using the same syringe with which it was instilled. All samples were immediately stored on wet ice and subsequently frozen at −80°C within 1 h of sample collection.

**Protein digestion and mass spectrometry analysis.** CVL samples were thawed on wet ice, vortexed, and centrifuged to pellet cellular debris. Supernatants free of cellular debris were used for all downstream processing. Portions (100 μg) of CVL protein, as determined by standard BCA protein assay (Novagen), from each participant were processed according to the protocol used by Birse et al. (27, 32). In brief, equal amounts of protein from each individual were digested with trypsin and analyzed using a label-free method. Analysis was performed on an LTQ Orbitrap XL mass spectrometer coupled to an Easy nLC II Nanoflow liquid chromatography system (Thermo Fisher Scientific). Raw data were processed by Progenesis (Nonlinear Dynamics) and Mascot (Matrix Science) using the Swiss-Prot human and bacterial protein database in 2013. A protein mix composed of all samples included in the study was injected every 10 samples throughout the length the experiment’s run time on the mass spectrometer, and these mixes were used to determine technical variability. Proteins that had a covariance of >25% across the mixes were excluded from further downstream analysis.

**Statistical and pathway analysis.** All protein abundances were normalized by the mean and log transformed (base 2). Protein abundance differences between individuals in their follicular phase versus those in their luteal phase were determined using unpaired, nonparametric Mann-Whitney U tests due to the study’s small sample size. Multiple comparisons correction with a false discovery rate of q < 0.05 only yielded three significant findings (P ≤ 0.0005). This significance threshold was likely too stringent for the study since it was underpowered. This view was supported by an examination of the P value frequencies, which showed that the frequencies were not uniformly spread as expected if the null hypothesis of no effect was true, suggesting that a true effect was likely observed and was therefore not likely the result of an inflated alpha (see Fig. S1 in the supplemental material). Based on this information, the statistical stringency was relaxed to P < 0.05 to avoid type II error, and subsequent downstream analysis was performed without multiple-comparison correction. Unsupervised cluster analysis was performed on proteins found to be differentially abundant between phases (P < 0.05; fold change, ≥2) using complete linkage and Spearman rank correlation as the distance metric. Differentially abundant proteins were also analyzed using DAVID (Database for Annotation, Visualization, and Integrated Discovery, v6.7) and Ingenuity Pathway Analysis software to determine the top biological processes and molecular functions associated with each menstrual cycle phase. Right-tailed Fisher exact tests were used to calculate the probability that the association between each protein in the data set and the biological function or pathway was random. Pathways with a minimum of at least two proteins associated and a P value of <0.05 were considered to be enriched, and biological processes (P < 0.05) assigned activation z-scores as determined using Ingenuity Pathway Analysis software were further analyzed.

**PLS-DA/lasso multivariate analysis.** The minimum set of proteins necessary to classify women based on menstrual cycle phase (luteal phase versus follicular phase) was determined using the least absolute shrinkage and selection operator (LASSO) method and partial least-squares discriminant analysis (PLS-DA). Both were implemented using Matlab software (MathWorks, Natick, MA). Briefly, for LASSO, K-fold cross-validation was used to determine the optimum value of the tuning parameter (“s”), such that the resulting model had the lowest possible mean squared error for prediction. The resulting features were chosen as the minimum set of biomarkers, and PLS-DA was used to assess the prediction ability of LASSO-selected biomarkers for classifying luteal- and follicular-phase groups. The data were normalized with mean centering and variance scaling prior to analysis, and cross-validation was performed by iteratively excluding random subsets (in groups of 18 data points) during model calibration and then using excluded data samples to test model predictions. This cross-validation ensured that the model was not over fit. In order to determine whether our LASSO-selected signature was significantly better for separating groups than other combinations of measured proteins, we also compared our LASSO-selected signature to 10,000 different combinations of the remaining non-LASSO proteins. For each combination, we computed classification and cross-validation error of the resulting PLS-DA model. We then computed an empirical P value to determine whether our LASSO-selected signature was significantly better than other protein combinations for differentiating women based on menstrual cycle phase.

**Gene set enrichment analysis (GSEA).** Normalized protein abundance values were uploaded to the Broad Institute’s gene set enrichment tool (http://www.broadinstitute.org/gsea) and compared against the curated Immunological Signatures gene set (C7.all.v4.0). Gene rank was calculated from the normalized protein abundance levels using the signal-to-noise metric. The permutation type was set to “phenotype,” gene set size parameters were set between 5 and 500 proteins associated, and data sets were not collapsed and were left in their original format. Only human gene sets with a nominal P value below 0.05 were included in our analysis. For gene sets with overlapping associations with our data set, only the top significantly enriched (lowest P value) gene set is shown. We defined normalized enrichment scores (NES) greater than an absolute value of 2.0 as high scoring, NES > |1.5| as medium scoring and NES < |1.5| as weak scoring associations.

**RESULTS**

**Participant characteristics.** Menstrual cycle phases of study participants (n = 19) were defined by days since last menstrual period where day 1 represented the first day of menses. The follicular phase included women (n = 7) who had samples collected on days 7 to 10 and samples collected from women on days 20 to 25 were assigned to the luteal phase (n = 12). There was no significant difference in clinical variables, including age (18 to 47 years; Mann-Whitney U test, P = 0.15) and condom use (Fisher exact test, P = 0.52), between the two groups. All individuals tested negative for STIs and bacterial vaginosis.

**Molecular pathways associated with immune cell recruitment are enriched during the luteal phase.** Mass spectrometry analysis of CVL samples collected from women in both the follicular and the luteal phases identified a total of 384 unique proteins, 19 of which were bacterial. Of these unique proteins, 43 (1 bacterial protein) were found to be differentially abundant between menstrual cycle phases (Mann-Whitney U test, P < 0.05, ≥2-fold change, Fig. 1). Hierarchical clustering of these significantly different proteins clearly distinguished women based on the menstrual cycle phase, with 26 (60%) and 17 (40%) proteins found at higher levels during the follicular and luteal phases, respectively (Fig. 2A). DAVID and Ingenuity Pathway Analysis were used to elucidate known biological functions associated with these phase-specific enriched proteins. The top biological processes positively associated with the follicular phase (positive activation z-scores below 2) were cell proliferation (z-score = 0.851, n = 16, P = 4.31E−5), protein metabolism (z-score = 1.199, n = 5, P = 1.71E−2), and migration of cells (z-score = 0.218, n = 7, P = 3.09E−2), whereas cell movement (z-score = −0.238, n = 8, P = 1.78E−2), apoptosis (z-score = −0.933, n = 10, P = 7.86E−3), cell death (z-score = −1.018, n = 12, P = 4.93E−3), and necrosis (z-score = −1.076, n = 9, P = 2.02E−2) were negatively associated with this time of the menstrual cycle (Fig. 2B, branch 2). Furthermore, the top molecular function associated with protein factors enriched during the follicular phase was endopeptidase inhibitor activity (n = 4, P = 1.5E−3), and the top canonical pathways included various cell growth and proliferation and met-
abiolic pathways (EIF2 signaling, n = 3, P = 1.47E–3; TR/RXR activation, n = 2, P = 4.87E–3; regulation of eIF4 and p70S6K signaling, n = 2, P = 1.38E–2; mTOR signaling, n = 2, P = 2.22E–2) and cell adhesion/junction pathways (remodeling of epithelial adherens junctions, n = 2, P = 3.15E–3; epithelial adherens junction signaling, n = 2, P = 1.38E–2; Fig. 2C) demonstrating follicular-phase enriched proteins’ involvement in important tissue regrowth and cell-cell adhesion processes within the epithelium. All of these biological functions are important for tissue growth, maintaining epithelial barrier integrity, and downregulating the inflammatory response. For example, cornifin B, cornulin, desmoglein 3, and cystatins are components of the epithelium that are important for generating and maintaining a cohesive keratinocyte barrier that is both flexible and yet impermeable (33–36). Although serine protease inhibitors (serpins) aid in wound healing, epithelial damage repair, and the inhibition of inflammatory proteases, all of these functions contribute to a healthy, adherent, defensive barrier (37–39).

Interestingly, the top biological processes predicted to be activated (Activation z-score ≥ 2) during the luteal phase included the following: the inflammatory response (z-score = 3.147, n = 12, P = 5.84E–7); the killing of bacteria (z-score = 2.76, n = 8, P = 1.11E–12); cell movement of leukocytes (z-score = 2.39, n = 7, P = 5.39E–5), including T lymphocytes (z-score = 2.129, n = 5, P = 1.87E–4); chemotaxis of leukocytes (z-score = 2.377, n = 8, P = 2.25E–6); degranulation of phagocytes (z-score = 2.2, n = 5, P = 2.42E–5); and damage of the epithelial tissue (z-score = 2.0, n = 4, P = 5.78E–5) (Fig. 2B, branch 1). The top molecular functions included endopeptidase activity (n = 6, P = 4.7E–5) and calcium ion binding (n = 7, P = 3.4E–4). Interleukin-8 (IL-8;
z-score = 2.0, n = 4, P = 1.41E–5) and leukocyte extravasation signaling (n = 3, P = 5.6E–4) were two of the top-scoring immune pathways associated with the luteal phase. These pathways are largely associated with the inflammatory response, with IL-8 having particular chemotactic effects on neutrophils. Other immune-related pathways associated with the luteal phase included immune cell trafficking (granulocyte adhesion and diapedesis, n = 2, P = 9.42E–3; agranulocyte adhesion and diapedesis, n = 2, P = 1.07E–2), the cellular immune response (production of nitric oxide and reactive oxygen species in macrophages, n = 2, P = 9.72E–3), and cellular movement (ILK signaling, n = 2, P = 1.04E–2) (Fig. 2C). These findings suggest that the luteal phase is a more immunologically active period than the follicular phase of the menstrual cycle.

A minimal multivariate protein signature that distinguishes the luteal phase includes cytoskeletal elements associated with tissue degradation/remodeling. Examining individual proteins in isolation most likely under-represents complex biological systems, where interconnected networks and relationships between various proteins are likely important for immunological processes. Here, we performed a multivariate analysis using data-driven modeling to better elucidate these interactions. Initially, we...
used the LASSO method for regression and shrinkage as a feature reduction method to our protein expression data set of 384 proteins. This algorithm identified the minimum set of proteins \( n = 18 \) that best differentiated the follicular phase from the luteal phase. This assessment was based on proteins that demonstrated phase-specific multicollinearity, which is due to similarities in covariance and/or biologically relevant relationships between proteins. We then used partial least-squares discriminant analysis to assess the predictive ability of these 18 proteins to generate new insight into relationships between these factors and menstrual cycle phase. Our model demonstrated 100% calibration accuracy, 93% cross-validation accuracy with the first two latent variables accounting for 45% of the variance \( P = 0.0013 \). Latent variable 1 (LV1) differentiated participants in the luteal phase (positive scores on LV1) from participants in the follicular phase (negative scores on LV1). Loading plot of the 18 LASSO-identified biomarkers. Eight of the eighteen identified markers were positively loaded on LV1, indicating that they were positively associated with the luteal phase or positively associated with the follicular phase.

**FIG 3** Multivariate model based on LASSO-selected biomarkers and partial least-squares discriminant analysis (PLSDA) classification. Our multivariate analysis expanded upon our univariate analysis by uncovering 10 new biologically relevant biomarkers that did not meet the univariate statistical thresholds. (A) Scores plot of the 19 individuals included in the present study demonstrating clear classification based on menstrual cycle phase with 100% calibration accuracy and 93% cross-validation accuracy, with the first two latent variables accounting for 45% of the variance \( P = 0.0013 \). Latent variable 1 (LV1) differentiated participants in the luteal phase (positive scores on LV1) from participants in the follicular phase (negative scores on LV1). (B) Loading plot of the 18 LASSO-identified biomarkers. Eight of the eighteen identified markers were positively loaded on LV1, indicating that they were positively associated with the luteal phase or positively associated with the follicular phase.

**TABLE 1** Three gene sets were positively associated with medium strength with the follicular phase using our selection criteria. The gene set enrichment analysis (GSEA) toolset was used to deduce phenotypic associations by comparing the results of the present study to independently generated immune cell gene expression patterns. Three gene sets were positively associated with medium strength (absolute normalized enrichment score [NES] > 1.5) with the luteal phase, and six gene sets were positively associated with medium strength with the follicular phase using our selection criteria.
TABLE 1 Proteins identified via LASSO-based multivariate modelinga

<table>
<thead>
<tr>
<th>Gene</th>
<th>Descriptiona</th>
<th>Species</th>
<th>Function(s)</th>
<th>Log, fold change (L − F)b</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARF4</td>
<td>ADP-ribosylation factor 4</td>
<td>Homo sapiens</td>
<td>Activation of phospholipase activity, vesicle-mediated transport</td>
<td>0.47</td>
<td>0.288</td>
</tr>
<tr>
<td>AZU1</td>
<td>Azurocidin*</td>
<td>Homo sapiens</td>
<td>Cellular extravasation, inflammatory response</td>
<td>3.33</td>
<td>0.001</td>
</tr>
<tr>
<td>CLCA4</td>
<td>Calcium-activated chloride channel regulator 4*</td>
<td>Homo sapiens</td>
<td>Chloride transport</td>
<td>−1.50</td>
<td>0.027</td>
</tr>
<tr>
<td>CLJC1</td>
<td>Chloride intracellular protein 1</td>
<td>Homo sapiens</td>
<td>Chloride transport, signal transduction</td>
<td>−0.76</td>
<td>0.021</td>
</tr>
<tr>
<td>SPRR1B</td>
<td>Cornifin B*</td>
<td>Homo sapiens</td>
<td>Epidermis development, keratinization</td>
<td>−3.18</td>
<td>0.001</td>
</tr>
<tr>
<td>DSG3</td>
<td>Desmoglein 3*</td>
<td>Homo sapiens</td>
<td>Cell adhesion</td>
<td>−1.11</td>
<td>0.012</td>
</tr>
<tr>
<td>EEF1B2</td>
<td>Elongation factor 1B*</td>
<td>Homo sapiens</td>
<td>Protein biosynthesis</td>
<td>−2.36</td>
<td>0.008</td>
</tr>
<tr>
<td>tuf</td>
<td>Elongation factor Tu*</td>
<td>Lactobacillus johnsonii</td>
<td>Protein biosynthesis</td>
<td>−2.36</td>
<td>0.045</td>
</tr>
<tr>
<td>KRT10</td>
<td>Keratin, type I cytoskeletal 10</td>
<td>Homo sapiens</td>
<td>Structural constituent of epidermis</td>
<td>0.82</td>
<td>0.138</td>
</tr>
<tr>
<td>KRT76</td>
<td>Keratin, type II cytoskeletal 2 oral*</td>
<td>Homo sapiens</td>
<td>Plays role in terminal cornification</td>
<td>1.56</td>
<td>0.060</td>
</tr>
<tr>
<td>KRT75</td>
<td>Keratin, type II cytoskeletal 75</td>
<td>Homo sapiens</td>
<td>Structural molecular activity</td>
<td>2.32</td>
<td>0.038</td>
</tr>
<tr>
<td>LMNB1</td>
<td>Lamin-B1</td>
<td>Homo sapiens</td>
<td>Apoptotic process</td>
<td>−0.30</td>
<td>0.620</td>
</tr>
<tr>
<td>S100A9</td>
<td>Protein S100-A9</td>
<td>Homo sapiens</td>
<td>Innate immune response</td>
<td>−0.89</td>
<td>0.124</td>
</tr>
<tr>
<td>GD12</td>
<td>Rab GDP dissociation inhibitor β*</td>
<td>Homo sapiens</td>
<td>GTPase activation</td>
<td>−1.55</td>
<td>0.0001</td>
</tr>
<tr>
<td>slpH</td>
<td>S-layer protein</td>
<td>Lactobacillus helveticus</td>
<td>Structural constituent of cell wall</td>
<td>1.33</td>
<td>0.169</td>
</tr>
<tr>
<td>WFD2C</td>
<td>WAP four-disulfide core domain protein 2</td>
<td>Homo sapiens</td>
<td>Broad range protease inhibitor</td>
<td>1.94</td>
<td>0.180</td>
</tr>
<tr>
<td>AZG1P</td>
<td>Zinc-α-2-glycoprotein</td>
<td>Homo sapiens</td>
<td>Negative regulation of cell proliferation, positive regulation of T cell-mediated cytotoxicity</td>
<td>0.99</td>
<td>0.331</td>
</tr>
</tbody>
</table>

a LASSO-based multivariate modeling accurately distinguishes between the luteal (L) and follicular (F) phases of the menstrual cycle.

b c, Overlapping proteins as determined by the Mann-Whitney U-test.

That is, the log, fold change difference between the luteal-phase abundance and the follicular-phase abundance.

outlined in Materials and Methods (see Table S2 in the supplemental material, P < 0.05). The top two follicular-phase-associated gene sets were unstimulated/unexposed, control immune cell types, including genes upregulated in macrophage controls relative to those exposed to Leishmania major (8 proteins, NES = −1.87, P < 0.001, Fig. 4C) and genes upregulated in monocyte controls relative to those treated with lipopolysaccharide (LPS; 12 proteins, NES = −1.76, P = 0.008, Fig. 4D). In contrast, the top two gene sets significantly enriched within the luteal phase were associated with neutrophils (6 proteins, NES = 1.55, P = 0.029, Fig. 4B) and HIV target cells (naïve CD4+ T cells; 7 proteins, NES = 1.57, P = 0.011, Fig. 4A). Although these independent data sets were not generated from mucosal samples, and acknowledging the limitations of comparing mucosal proteomic data to that of cellular genomic data sets, these data do suggest that there are immunological differences between the luteal phase and follicular phase, a view which agrees with the preceding pathway analysis.

DISCUSSION

This study is the first comprehensive proteomic analysis of the changes in the mucosal immunological environment that exists between the phases of the menstrual cycle in humans. These data indicate there are significant differences between the female genital tract mucosal proteome during the follicular and luteal phases of the menstrual cycle and generates new hypotheses about potential mechanisms of HIV susceptibility (Fig. 5). Factors overabundant during the follicular phase were associated with biofunctions that underlie cellular adherence and epithelial barrier integrity, as well as inflammation mediation (39–41). This may help in maintaining an environment less likely to contribute to HIV infection since various studies have highlighted the role of inflammatory conditions such as STIs and some forms of bacterial vaginosis in increasing susceptibility to HIV infection (42, 43). Protease inhibitors were found at higher levels during the follicular phase and are known regulators of inflammation. Higher levels of these antiproteases may help to reduce inflammatory processes, while at the same time perform innate antiviral functions (39), which may in turn help prevent the establishment of an HIV founder population if a woman is infected during this time. In fact, higher levels of antiproteases have been associated with HIV resistance in highly exposed seronegative female sex workers (23, 44). Furthermore, the anti-inflammatory effects of these factors may also promote the maintenance of the epithelial barrier throughout the genital tract, which is also critical for defense against pathogen infection, including HIV (45). Indeed, our study also found proteins associated with barrier integrity to be overabundant during the follicular phase.

In contrast, the luteal phase was primarily associated with factors involved with chemotaxis, cellular movement, and immune cell activation, including that of HIV target cells, as well as peptidase activity and damage to the epithelial barrier. These biofunctions have all been associated with enhanced HIV infection and/or transmission (20, 46). Given the current understanding of the late luteal phase and the inflammatory processes associated with endometrial bleeding (47), these findings agree with the predicted physiological events that may occur during this time. The withdrawal of progesterone during the late luteal phase is known to trigger a signaling cascade that leads to the beginnings of tissue remodeling and the infiltration of leukocytes, including neutrophils, macrophages, and natural killer cells. The signatures of these biological processes are reflected in our analyses, including the IL-8 and leukocyte extravasation signaling pathways, an observation that agrees with what is known to occur as the body prepares for the onset of menstruation (47). These findings also overlap the findings of a separate study our group conducted in
FIG 4 Top two gene sets significantly enriched during the luteal phase (A and B) and the follicular phase (C and D) based on GSEA. Gene sets associated with the luteal phase represented specific immune cell signatures for neutrophils and CD4+ T lymphocytes, whereas gene sets associated with the follicular phase represented unstimulated/unexposed, control immune cell phenotypes. The gene sets were as follows. (A) Genes upregulated in naive CD4 T cells relative to naive B cells (7 proteins, NES = 1.57, \( P = 0.011 \)). (B) Genes upregulated in neutrophils relative to basophils (6 proteins, NES = 1.55, \( P = 0.029 \)). (C) Genes upregulated in macrophage controls relative to \( \text{Leishmania major} \) exposed macrophages (8 proteins, NES = 1.87, \( P = 0.001 \)). (D) Genes upregulated in monocyte controls relative to LPS treated monocytes (12 proteins, NES = 1.76, \( P = 0.008 \)).
pig-tailed macaques in which many of the same factors (CTSG, MMP9, MPO, MMP8, and LCP1) involved in tissue permeability and leukocyte extravasation signaling were found to be overabundant during the macaque luteal phase (48). This provides further evidence that similar molecular events are occurring during the luteal phase in both macaques and humans and represent potential pathways contributing to enhanced HIV susceptibility.

Further to our univariate analysis, our multivariate model uncovered a molecular pattern that may be more reflective of a global response occurring over the course of the menstrual cycle. This is not surprising given the difficulty in determining true biological effects from probabilistic noise since many of the changes that occur in vivo are of a small magnitude that would not pass univariate statistical thresholds. This analysis identified positive associations between cytoskeletal elements and proteases/chemoattractants such as azurocidin during the luteal phase, as well as negative associations with cell-cell adhesion and cornification. These inferred biological events may be involved in tissue remodeling, which must occur prior to leukocyte infiltration (49). This time of potentially increased epithelial barrier disruption may be a key feature contributing to HIV susceptibility since both in vivo macaque models and ex vivo human explant tissues show enhanced viral diffusion when the cellular junctions were compromised (16). Furthermore, the use of progesterone-based intruterine devices has been associated with ectocervical epithelial thinning and decreased mRNA levels of important tight-junction proteins (50). This provides further evidence of the impact that progesterone levels can have on the physical barrier of the genital tract whether expressed endogenously during the luteal phase or through exogenous application.

Immune cell numbers are also known to be influenced by the menstrual cycle and hormone levels. For instance, neutrophils are known to increase in number in the endometrium during the late luteal phase, as well as during menses following an IL-8 surge (13, 47). This influx is believed to occur for the purposes of aiding in endometrial tissue breakdown via the release of proteases and to increase immune defense while the epithelial barrier is disrupted. Some studies have noted CD4+ T cell population differences during the menstrual cycle with increased levels during menstruation (51), while others did not find increased CD4+ T cell numbers in cervical explant tissues collected from the luteal phase or in progesterone-based contraceptive users (10, 50). Although cell populations were not assessed in our study, our data suggested that neutrophil and CD4+ T cell signatures were associated with the luteal phase by GSEA. A recent study by Arnold et al. found that elevated proteases (MMP9, MMP8, and neutrophil elastase) and cytoskeletal elements, a signature very similar to the luteal phase described here, positively associated with increased numbers of cervical CD4+ T cells collected by cytobrush. This strongly suggests there may be a link between CD4+ T cell numbers and mucosal protease levels (30). Interestingly, one protease in particular, azurocidin (AZU1) represented the highest ranking factor associated with the CD4+ T cell gene set enriched during the luteal phase based on our GSEA. Furthermore, this same factor was also the most strongly correlated luteal-phase biomarker from our multivariate model. AZU1’s associative properties with the extracellular matrix and its function as an antimicrobial and inflammatory mediator with roles in chemotaxis and cellular extravasation readily suggests that it may play a critical role in the HIV target cell recruitment predicted to occur during the luteal phase (52, 53). Other neutrophil granule-derived proteins such as cathelicidin and alpha-defensins, which were also enriched during the luteal phase, have been associated with enhanced HIV acquisition risk when found at high levels in the genital secretions of both men and women (21, 22). Indeed, the shared chemotactic activities of these neutrophil granule-derived proteins could be the driving force behind the immune cell recruitment and the corresponding epithelial barrier disruption observed during the luteal phase in the present study, which may contribute to the overall vulnerability to HIV infection observed during this time. It is possible that the HIV target cell populations and/or phenotypes, which can impact HIV acquisition risk due to an increased expression of HIV coreceptors (54), varied at the mucosal surface of these luteal-phase samples; however, we were unable to confirm this in our study since matching cytobrush and/or biopsy samples were not collected for this purpose. A longitudinal study assessing these immune cell populations in parallel with these molecular signatures would help answer this question.

A hormone-mediated effect on HIV target cell influx into the female genital tract has been observed in the cervices of young women who use the progesterin-based injectable, depot medroxyprogesterone acetate (DMPA). This preliminary study associated DMPA use with a 5-fold increased risk of HIV acquisition compared to noninjectable user counterparts (55). Our findings indicating a general immune activation phenotype during the luteal phase may reflect a response similar to that induced by injectable contraceptives. For instance, a study examining DMPA use and a list of predefined innate immune factors found that alpha-defensins 1 to 3, LL-37 (cathelicidin), and lactoferrin were found at significantly higher levels in the secretions of DMPA users (56). Indeed, these findings match what was found in our study during the progesterone-dominant luteal phase. Exogenous progesterin is known to be rapidly degraded shortly after its injection (57), which may mimic the rise and fall of progesterone during the luteal phase. With this consideration in mind and these studies’ findings, future studies examining the mechanism of progesterin-mediated HIV susceptibility should pay special attention to when study participants receive their hormonal injections.

Our study had a few limitations. Since the study was designed in a post hoc nature, we were unable to confirm or ascertain immune cell populations and other biological events in our analysis; these are important next steps in ongoing studies. Furthermore, other caveats include the study’s small sample size, the fact that it...
was performed cross-sectionally, utilizing days since last menstrual period as an indicator of menstrual cycle phase, and the lack of systemic hormone measurements. The cross-sectional nature of the study represents a major caveat since we were unable to compare phase-specific changes occurring within the same women, which would be evident in a longitudinal study. Hormonal measurements would help to better elucidate phases since not all women have matching cycles. However, to mitigate this issue, we chose to set very narrow windows for our follicular- and luteal-phase inclusion criteria to prevent phase mismatch or overlap. Our study did not find phase-specific changes in bacterial species expression at the mucosa of the lower female genital tract. Specific *Lactobacillus* proteins were preferentially enriched in one phase or the other, but no biological effect could be deduced from this finding. Our findings are only inclusive of the 19 bacterial proteins confidently identified using our proteomics method; however, our results appear to be consistent with those of a pyrosequencing study examining the microbiota changes over the menstrual cycle of Canadian women. The study found the vaginal microbiome of most women to be relatively stable, and most women were predominantly colonized by *Lactobacillus* and *Bifidobacterium* species (58).

Our study elucidates the immunological events that occur during the luteal phase that may increase women’s susceptibility to HIV infection, as has been observed in studies of SHIV-challenged nonhuman primates and HIV-challenged human cervical explant models (8–10). This includes immune cell recruitment and activation processes, as well as tissue remodeling and barrier disruption. Furthermore, it is possible that these same processes may be induced or enhanced through the use of progesterone-based hormonal contraceptives and should be examined. Overall, these findings indicate that an inflammatory and immune-activated state is prevalent in the female genital tract during the progesterone-dominated, mid to late luteal phase. This knowledge aids in the generation of new hypotheses regarding the potential mechanisms of hormone-related susceptibility to HIV.

**ACKNOWLEDGMENTS**

This study was funded by Canadian Institutes of Health Research grant HBP137699 (A.B.), the Public Health Agency of Canada (T.B.B. and G.R.W.), Institute for Collaborative Biotechnologies of the U.S. Army Research Office grant W911NF-09-0001 (K.B.A. and D.A.L.), and a gift from the Begg Family to the MIT Center for Gynepathology Research Office grant W911NF-09-0001 (K.B.A. and D.A.L.), and a gift from the Begg Family to the MIT Center for Gynepathology Research Office grant W911NF-09-0001 (K.B.A. and D.A.L.), and a gift from the Begg Family to the MIT Center for Gynepathology Research Office grant W911NF-09-0001 (K.B.A. and D.A.L.). We thank all of the study participants. We thank Fiona Lynch, Ellen Verlen, and Regina Harden for collecting the samples and providing epidemiological data. We also thank Max Abou for technical support and Laura Romas and Dan Chatue for assisting with the statistical analysis.

**REFERENCES**

1. UNAIDS. 2012. Women out loud: how women living with HIV will help to end AIDS. UNAIDS, Geneva, Switzerland.


September 2015 Volume 89 Number 17 Journal of Virology jvi.asm.org 8803

Downloaded from http://jvi.asm.org on September 29, 2015 by UNIV OF ILLINOIS AT CHICAGO


