Title

Molecular Classification of Endometriosis and Disease Stage Using High-Dimensional Genomic Data

Journal

Endocrinology, 155(12)

ISSN

0888-8809

Authors

Tamaresis, John S
Irwin, Juan C
Goldfien, Gabriel A
et al.

Publication Date

2014-12-01

DOI

10.1210/en.2014-1490

Peer reviewed

Endocrinology. 2014 Dec; 155(12): 4986–4999. PMCID: PMC4239429
Published online 2014 Sep 22. doi: 10.1210/en.2014-1490

Molecular Classification of Endometriosis and Disease Stage Using High-Dimensional Genomic Data

John S. Tamaresis, Juan C. Irwin, Gabriel A. Goldfien, Joseph T. Rabban, Richard O. Burney, Camran Nezhat, Louis V. DePaolo, and Linda C. Giudice

Center for Reproductive Sciences (J.S.T., J.C.I., G.A.G., L.C.G.), Department of Obstetrics, Gynecology and Reproductive Sciences, and Department of Pathology (J.T.R.), University of California, San Francisco, California 94143; Department of Obstetrics and Gynecology and Clinical Investigation (R.O.B.), Madigan Healthcare System, Tacoma, Washington 98431; Department of Obstetrics and Gynecology (C.N.), Stanford University, Stanford, California 94024; and Fertility and Infertility Branch (L.V.D.), *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892

Corresponding author.

Address all correspondence and requests for reprints to: Linda C. Giudice, MD, PhD, Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, 505 Parnassus Avenue, Room 1496, San Francisco, CA 94143-0132., E-mail: giudice@obgyn.ucsf.edu.

Received 2014 Jun 17; Accepted 2014 Sep 15.

Abstract

Endometriosis (E), an estrogen-dependent, progesterone-resistant, inflammatory disorder, affects 10% of reproductive-age women. It is diagnosed and staged at surgery, resulting in an 11-year latency from symptom onset to diagnosis, underscoring the need for less invasive, less expensive approaches. Because the uterine lining (endometrium) in women with E has altered molecular profiles, we tested whether molecular classification of this tissue can distinguish and stage disease. We developed classifiers using genomic data from n = 148 archived endometrial samples from women with E or without E (normal controls or with other common uterine/pelvic pathologies) across the menstrual cycle and evaluated their performance on independent sample sets. Classifiers were trained separately on samples in specific hormonal milieu, using margin tree classification, and accuracies were scored on independent validation samples. Classification of samples from women with E or no E involved 2 binary decisions, each based on expression of specific genes. These first distinguished presence or absence of uterine/pelvic pathology and then no E from E, with the latter further classified according to severity (minimal/mild or moderate/severe). Best performing classifiers identified E with 90%–100% accuracy, were cycle phase-specific or independent, and used relatively few genes to determine disease and severity. Differential gene expression and pathway analyses revealed immune activation, altered steroid and thyroid hormone signaling/metabolism, and growth factor signaling in endometrium of women with E. Similar findings were observed with other disorders vs controls. Thus, classifier analysis of genomic data from endometrium can detect and stage pelvic E with high accuracy, dependent or independent of hormonal milieu. We propose that limited classifier candidate genes are of high value in developing diagnostics and identifying therapeutic targets. Discovery of endometrial molecular differences in the presence of E and other uterine/pelvic pathologies raises the broader biological question of their impact on the steroid hormone response and normal functions of this tissue.

Endometriosis (E) is a progressive, debilitating, estrogen-dependent, progesterone (P_4)-resistant, inflammatory disorder associated with pelvic pain and infertility, with endometrium (uterine lining)-like tissue present outside the uterus (1, 2). By retrograde menstruation, endometrial tissue fragments/cells are transplanted to the pelvis (3), where they establish a blood supply, respond to cyclic hormones, grow, invade surrounding structures, become innervated (4, 5), and elicit a local inflammatory response and scarring (1, 2). E affects 6%–10% of reproductive-age women (6) and 50% of women with pelvic pain and/or infertility (>100 million women worldwide) (7) and is a major cause of disability and compromised quality of life (8, 9). Pelvic, lower abdominal and back pain, and urinary and gastrointestinal symptoms make diagnosis challenging, because many symptoms are nonspecific or are associated with other disorders (1). Pelvic inflammation and nerve infiltration result in pain (4, 5), and infertility is due to ovulatory dysfunction, poor egg quality, abnormal (P_4-resistant) uterine endometrium, and compromised embryo implantation (1, 2, 10). In 2009, estimated United States healthcare costs for diagnosis and treatment of E-related pain and infertility totaled $49 billion (9).

The current gold standard for E diagnosis and staging is surgical visualization under general anesthesia with histologic confirmation of endometrial glands and stroma in biopsied pelvic lesions (1). Drawbacks of surgical diagnosis include procedural and anesthetic risks, time away from work and family, and cost (1, 9). The mean time from pain onset to surgical diagnosis is 6.7–11.0 years (8, 11), with attendant risk of disease progression (12) over this interval. Diagnostic delay may have deleterious consequences (11, 13), including progression of pain and infertility requiring more aggressive treatment approaches. In addition, recent data suggest radical removal of all visible disease is protective against later developing ovarian cancer (14), underscoring the importance of early diagnosis and intervention. When E is suspected, pain is empirically treated with contraceptive steroids, antiinflammatory drugs, and GnRH agonists, but these therapies are unsatisfactory in 20% of women because of side effects or resistance (1). Thus, a prompt, low-risk, low-cost diagnostic with high accuracy is needed (15) to shorten time to diagnosis, minimize disease progression and ovarian cancer risk, optimize timing and strategies for pain and infertility therapies, and monitor disease recurrence.

Because the endometrial transcriptome differs significantly in women with vs without disease (16,–20), herein we applied machine learning and high dimensional analysis to leverage these observed differences towards developing classifiers for E diagnosis and stage and pursue further insights into the pathobiology of endometrium in women with disease. We report highly accurate diagnostic classifiers that use sequential binary decisions, each based on specific gene sets, that distinguish E (disease and stage) and are menstrual cycle stage specific or hormonal milieu independent. Furthermore, differential gene expression and pathway analyses, based on these binary decisions, provided insight into steroid hormone signaling and other molecular and cellular dysfunctions in endometrium of women with E and also dysfunctions with other uterine/pelvic disorders.

Materials and Methods

Specimens and processing

In the National Institutes of Health (NIH)/University of California, San Francisco (UCSF) Human Endometrial Tissue Bank, n = 148 archived endometrial samples from different menstrual cycle phases were identified from women with E and pelvic pain and/or infertility (n = 77), no E but with uterine/pelvic pathology (NE.UPP) (n = 37, including symptomatic uterine fibroids, pelvic organ prolapse, and adenomyosis), and n = 34 no E and no uterine/pelvic pathology (NE.NUPP) (normal controls). Included were cycling women 20–50 years old undergoing procedures for diagnosis and treatment of pelvic pain, infertility, or benign gynecologic conditions and normal volunteers (Supplemental Table 1). Exclusion criteria included hormonal treatment within previous 3 months and presence of malignancy or major systemic disease. Uterine/pelvic abnormalities were verified from participants' operative and pathology

reports, and E stage was per the revised American Fertility Society classification system (21). Menstrual cycle phase was assigned by endometrial histology (22), reviewed by 2 pathologists, confirmed by serum estradiol and P_4 levels (NIH SCCPIR Ligand Core, University of Virginia, Charlottesville, VA), and corroborated by 2 independent bioinformatics methods: clustering in unsupervised whole-transcriptome principal component analysis (23) and cycle phase assignment classifier analysis. Samples were identified as proliferative phase endometrium (PE), early-secretory phase endometrium (ESE), or midsecretory phase endometrium (MSE). Four samples with conflicting ("indeterminate") cycle phase assignment were included only for development of cycle phase-unrestricted classifiers (see below). All samples were obtained after written, informed consent through the UCSF Committee on Human Research Protocol 10-02786.

Whole-tissue samples were processed using rigorous protocols and hybridized to whole-genome microarrays, as described (17, 23). Total RNA was purified from specimens, as described (23), and only high-quality RNA samples (RNA integrity number >7) were processed for hybridization to Affymetrix HU133 Plus 2.0 (54 675 probe sets) high-density oligonucleotide arrays at the UCSF Genomics Core Facility.

Data processing and statistical analysis

Classifier development The .cel data files were uploaded to the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) Database (series accession number GSE51981). All data analyses were performed using R and Bioconductor. Microarray data for all 148 samples were simultaneously normalized using the Bioconductor package *GCRMA* (24). Classifier construction and validation were performed using margin tree classification (25) with the Bioconductor *marginTree* package. Different classifiers were developed with this methodology, using samples in a particular menstrual cycle phase (phase specific), across 2 phases (phase restricted), or regardless of phase (phase unrestricted). For each classifier, corresponding samples were partitioned using stratified random sampling into 2 discrete subsets: 80% used for construction and 20% used for independent validation (see Figure 1 A). The number of samples in each disease category defined the stratification, thereby preserving the original proportional representation in both subsets. Classifier construction involved building the margin tree, followed by classifier training through k-fold cross-validation to minimize overfitting by further partitioning the construction set into k nonoverlapping folds (typically $k = 5$ to 10 folds), each preserving the proportional stratification of the original subset, then combining k-1 folds into a train set to build a classifier with the remaining fold as test set to score accuracy. This process was repeated until each fold had been used once as the test set, then the optimal value of the margin tree's adjustable parameter was found, creating a classifier that best generalized to new samples. Finally, performance of the developed classifier was evaluated on an independent validation sample set to compute the classification accuracy on samples never seen by the classifier during construction.

Resampling

Resampling (multiple iterations of random partitioning and classifier construction/validation) used a sequence of different prime numbers as seeds to initialize random partitioning of samples into construction and validation sets. This generated 250 classifiers per sample set, each one scored for performance on the corresponding independent validation set. This resampling approach added robustness, addressed the inherent bias introduced by partitioning of the sample sets, and allowed estimating validation accuracy distribution for the classifiers, how frequently a gene would be used, and its ranking in importance, for a specific binary decision (26).

Differential gene expression and pathway analyses

Differential gene expression analysis used R/Bioconductor package, *limma*, and moderated *t* statistic with false discovery rate error control (27). Comparisons were between samples grouped according to the classifier-defined decision tree branches and between E and no E but other abnormal uterine/pelvic pathology vs normal controls, respectively. Each contrast was evaluated simultaneously for each of the 3 phases. Biological functions and canonical pathway analyses for differential expression comparisons were conducted using R/Bioconductor package, *sigPathway* (28), including annotations from Gene Ontology, KEGG, Biocarta, and other sources. Pathways with changes in function were identified by aggregating differentially expressed member genes into a pair of pathway-level statistics, assessing the statistical significance of each, and ordering the pathways by the average rank of the 2 statistics. Activated or inhibited pathways were identified when both statistics were positive or negative, respectively.

Q-RT-PCR validation of gene expression

Ninety-two differentially expressed genes were analyzed using 48 proliferative endometrium samples (13 NE.NUPP, 13 NE.UPP, 11 E.Min/Mild, and 11 E.Mod/Severe) by Q-RT-PCR and the Fluidigm 96.96 Dynamic Array Integrated Fluidic Circuits and Biomark System with primers from Fluidigm Corp (Supplemental Table 2) (29). Briefly, cDNA samples were preamplified to enrich loci of interest using primer pair (500nM) pools for all genes assayed, diluted 1:20, and quantitative PCR performed using EvaGreen binding dye. Reaction conditions and amplicon integrity assessment were as described (29). Data were processed with individually set thresholds for each gene, and linear baseline correction using BioMark real-time PCR Analysis software (version 3.0.4). Metastasis associated lung adenocarcinoma transcript 1 (MALAT1) expression was assessed by Q-RT-PCR on Mx 3005 Pro (Stratagene), with 10-ng cDNA amplified with 150nM primer mix (Eurofins MWG Operon) and SensiFAST SYBR Lo-ROX (Bioline) for 40 cycles (29). Ribosomal protein L19, with stable endometrial expression by microarray and Q-RT-PCR, was used for normalization. Data were analyzed using Microsoft Excel office (2011) by the comparative Cycle threshold (Ct) method (30) to assess relative expression between groups. Dixon's *Q* test was used to identify and remove outliers (31).

Correlation between microarray and Q-RT-PCR

The agreement between microarray and Q-RT-PCR differential gene expression data was evaluated with nonparametric Spearman's Rho and Kendall's Tau using StatView 5.0.1 (SAS Institute, Inc). For both tests, a positive value indicates agreement between the microarray and Q-RT-PCR data at the significance level given by the *P* value assessed with a two-tailed null hypothesis of no association.

Results

Diagnostic classifiers

Clinical data are shown in Table 1 and Supplemental Table 1. Gene expression analysis data were characterized by extreme asymmetry with many more variables (54 675 probe sets) than observations (148 samples), and the experimental design presented a multiclass problem with 3 disease categories for discrimination (NE.NUPP, NE.UPP, and E). Margin tree classification (25) effectively treated these experimental design and dataset features, whereas analytical alternatives (32, 33) yielded classifiers of low-fair accuracy. The margin tree algorithm used microarray data in an unsupervised manner to resolve the classification of more than 2 classes into a tree-like sequence of binary decisions each based on expression levels of distinct sets of genes. For each classifier, corresponding samples were partitioned into 80% construction and 20% independent validation sets (Figure 1A), and classifiers were developed with construction sets, underwent *k*-fold cross-validation to minimize overfitting, and were scored for accuracy with validation sets not used during construction.

Decision tree

Highest accuracies were obtained with composite classifiers comprising a disease component and a severity component, corresponding to the decision tree shown in Figure 1B. The disease component identifying endometrial samples from women with E included 2 sequential binary decisions. The first was whether pathology was absent (normal [NE.NUPP]) or present. If present, samples went to the next decision level and were classified as coming from women with NE.UPP or E. The severity component assigned disease stage and classified samples from women with E into minimal/mild (E.Min/Mild) or moderate/severe (E.Mod/Severe). This sequence of decisions was consistent across menstrual cycle phases and resampling iterations, implying robustness of the classification process.

Disease component

Disease classifiers discriminated among NE.NUPP, NE.UPP, and E. We developed menstrual cycle phase-unrestricted classifiers diagnosing samples in all cycle phase categories, phase-restricted classifiers diagnosing samples in both PE and ESE, and phase-specific classifiers (PE, ESE, or MSE) diagnosing samples in the corresponding cycle phase. Resampling yielded multiple high (>90%) accuracy classifiers, the best performing being the PE/ESE phase restricted, and all phase-specific classifiers achieving 100% accuracy on independent validation samples. Altogether, 75 high accuracy disease classifiers were discovered (Table 2). Characteristics of individual disease classifiers (accuracy and number of probe sets for each decision) from all diagnostic variants are in Supplemental Table 3A, wherein 2 particular PE and ESE phase-specific disease classifiers achieved 100% accuracy using less than 100 genes for each disease classification decision.

Severity component

Disease classifiers segregated E samples into 1 class regardless of severity; then, severity classifiers, constructed and validated with samples from women with E, divided them into minimal/mild and moderate/severe. We developed numerous phase-specific severity classifiers that achieved 100% accuracy on independent validation samples (Table 3). Four ESE severity classifiers used less than 100 genes, whereas the lowest number of probe sets used by a PE severity classifier was 196, and all MSE severity classifiers used more than 1000 probe sets (Supplemental Table 3B).

Core genes

Each binary decision in the classification process was based on expression levels of distinct gene sets specific for a given binary decision and classifier. For a given sample set, resampling generated multiple classifiers, and those performing with the same validation accuracy comprised a “classifier family,” and for a given binary decision, a defined subset of “core” genes was used by every classifier within a family. Thus, each classifier family has 3 distinct sets of core genes, 1 for each binary decision. Core genes, therefore, are indispensable for diagnosis at the defined accuracy by all classifiers within a given family. Core genes for the PE and ESE phase-specific/100% accuracy disease and severity classifier families are in Supplemental Table 4. For MSE, highest accuracies were 91% for disease and 100% for severity, and classifiers required many more core genes at all decision points (Supplemental Table 4).

Differentially expressed genes and pathway analyses

Classifier analysis defined a hierarchy of stratification for the disease categories providing a roadmap for further data mining to unveil underlying biological processes. The sequence of binary decisions consistently first segregated all pathologies combined from normal controls, indicating this represents the most salient transcriptomic contrast and implying relatedness of the transcriptomic signatures when pathology is present vis-à-vis a normal pelvis. Margin tree dendrograms (Figure 2A) confirmed that separation between the 2 pathology classes (E and NE.UPP) was substantially narrower than that between normal controls (NE.NUPP) and the combined pathology classes ([E + E.UPP]).

The core genes in Supplemental Table 4 correspond to the 3 decision levels (Figure 1B) for the highest accuracy phase-specific composite classifiers. Importantly, each level reflects margin tree algorithm decisions for optimal high-dimensional partition of the classes into two groups at each juncture (32), not pairwise comparisons of differentially expressed genes, and the classifier gene lists at each decision were not derived from differential expression analysis. Indeed, classifier-generated gene lists may not contain all differentially expressed genes between the 2 groups at a given decision level, and conversely, not all genes used by the classifier for a decision will be necessarily differentially expressed between the 2 groups at a conventional significance and fold change threshold. To determine genes differentially expressed in endometrial samples in the groups corresponding to the classifier branches, ie, [E+NE.UPP] vs controls (NE.NUPP), E vs NE.UPP, E.Mod/Severe vs E.Min/Mild, we conducted differential expression analyses using all samples for classifier development across 3 phases of the menstrual cycle. We also compared gene expression in E vs normal and NE.UPP vs normal and conducted pathway analyses and validated gene expression by Q-RT/PCR.

Genes differentially expressed in each of the comparisons (Supplemental Table 5) were uniquely differentially expressed in different hormonal milieu, in 1 phase of the cycle, in 2 phases, or across all 3 phases. Figure 2B shows the Venn diagram of differentially expressed genes for comparisons at each branch point in Figure 1B decomposed by phase into the subgroups shown, with overlaps in different phases. Consistent with classifier analysis, the greatest number of differentially expressed genes (21 946 probe sets) was found in [E+NE.UPP] vs NE.NUPP (Figure 2B), compared with only 214 in E vs NE.UPP, and 8444 in the severity comparison

Tanscriptomic changes in E and other pelvic/uterine pathologies

Branch 1

Differentially expressed genes across the cycle. Relation to highest performing classifier core genes and endometrial dysfunction.

In branch 1, the most highly differentially expressed genes across all phases studied included FOSB, FOS, early growth response 1, JUNB, deiodinase iodothyronine type II (DIO2), and integrins (Table 4). Striking were the high fold changes of FOSB and FOS and that they were part of the highest performing classifier core genes in the first binary decision in the proliferative and early-secretory phases (Supplemental Table 4A), distinguishing a normal pelvis (NE.NUPP) from abnormal [E+NE.UPP]. Twenty-five out of 88 core genes listed in Supplemental Table 4 were differentially expressed (>1.5-fold) in 1, 2, or 3 phases of the cycle.

Also of interest across all phases is up-regulation of genes involved in immune cell activation and signaling. Among down-regulated genes are members of the EGF signaling pathway known to be dysregulated in endometrium of women with E (16, 17, 34), including ERBB2 and ERBB2-interacting protein. Most genes were similarly regulated in E vs NE.NUPP and NE.UPP vs NE.NUPP, supporting few genes differentially expressed in all 3 phases in E vs NE.UPP (Figure 2B, Tables 5 and 6, and Supplemental Table 5).

Differentially expressed genes specific to each cycle phase. Some genes were up-/down-regulated in [E+NE.UPP] vs NE.NUPP with differential expression restricted to 1 phase of the cycle (Table 5; for complete list, see Supplemental Table 5). In proliferative phase, most highly up-regulated genes included glutathione peroxidase (GPX)3, water and ion channels, secretoglobin family members, and highly down-regulated genes included carboxypeptidase M, Wnt inhibitors, integrins, and matrix degrading enzymes. Many of these were also differentially expressed in E vs NE.NUPP and NE.UPP vs NE.NUPP (Supplemental Table 5). In early-secretory phase (Table 5), highest up-regulation was for immune modulators, transmembrane proteins, collagens, the cell cycle regulator PTEN, and transcription factors, and down-regulation of solute carriers, thyroid hormone signaling, phosphodiesterases, and ephrin family members. Midsecretory phase was remarkable for up-regulation of genes for immune response and

angiogenesis. Down-regulated genes involved ribosomal modification proteins, the immune response, prostaglandin synthesis, metalloproteinases, integrins, mucins, and nuclear receptor coactivator 2 (NCOA2). Most of these were also differentially expressed in E and NE.UPP vs normal, including genes for the killer immunoglobulin receptors (KIRs), with additional up-regulation of hemoglobin-β, IL family members, and hepatocyte growth factor (HGF), and down-regulation of matrix metalloproteinase 26 (MMP26), microRNA (MIR)30C2 in E, and up-regulation of HLA-DRB1, gastrin, secreted frizzled-related protein 1 (SFRP1), and down-regulation of solute carrier proteins, TGFB2, among others in NE.UPP (Supplemental Table 5).

Branch 2 There were remarkably few up-regulated probe sets in the proliferative phase in E vs NE.UPP (Table 6 and Supplemental Table 5). However, striking was down-regulation of the angiogenesis inducer, CYR61 (ie, up-regulated in NE.UPP), suggesting an even greater proangiogenic environment in endometrium of women with uterine fibroids and other uterine/pelvic disorders compared with E. The early-secretory phase was marked by up-regulation of genes for transmembrane proteins, stromal-derived factor 2, and down-regulation of a variety of transcription factors, ubiquitin pathway members, and others. The midsecretory phase displayed up-regulation of genes for solute carrier proteins, cytochrome P450 members, transcription factors, angiogenesis, GPCRs, and calpain. Down-regulated genes included those for junctional complexes, thyroid hormone action, mucins, and helicases. There were few genes commonly regulated in all phases of the cycle (Supplemental Table 5).

Pathway analyses in endometrium in the context of uterine/pelvic disorders

The most highly up and down activated canonical pathways are shown in Table 7 (see Supplemental Tables 6–10 for the complete set of *sigPathways* at http://obgyn.ucsf.edu/crs/research/giudice_lab/supplemental_tables_6–10.aspx. Striking in branch 1 comparison of [E+NE.UPP] vs normal (NE.NUPP) in proliferative phase (Table 7 and Supplemental Table 6A) were global down-regulation of cell division and activation of neuropeptide signaling, immune activation and steroid hormone metabolism. Most of these were also observed in E vs normal (Table 7 and Supplemental Table 9) and NE.UPP vs normal (Table 7 and Supplemental Table 10). Early-secretory phase was marked by activation of growth factor, steroid hormone, and integrin signaling pathways and immune activation in [E + NE.UPP] vs normal. Androgen receptor and ER signaling and angiogenesis were more prevalent in samples from women with E (Table 7 and Supplemental Table 6B). In midsecretory phase, cell defense and immune activation prevailed in [E+NE.UPP] vs normal (Table 7 and Supplemental Table 6C) with immune activation and stress/inflammatory response additionally in E vs normal, and the response to stress/toxicity prevailing in NE.UPP vs normal controls. Down-regulation of specific signaling pathways was observed uniquely in endometrium from women with these disorders vs controls (Table 7). In E (all stages) vs NE.NUP in early and midsecretory phases, aberrant cell signaling involving EGF, ERK/MAPK, and Jak/STAT/PI3K/protein kinase B (AKT) signaling was observed, consistent with previous studies in women with moderate/severe (16) or minimal/mild disease (17) vs controls, and a baboon model of E in midsecretory phase (35). Novel in the current study is the prominence of immune activation in E vs NE.NUP in all 3 phases, as well as transcriptomic analysis of endometrium from women with other uterine/pelvic pathologies and phenotypically normal controls.

For branch 2 (E vs NE.UPP) (Table 7 and Supplemental Table 7), there were only down-regulated pathways in the proliferative phase, in contrast to early-secretory phase, with only pathway activation. In the midsecretory phase, activated pathways in E involved response to chemical substance, ER and Wnt pathway signaling, cytoskeleton, and angiogenesis, and down-regulation involved DNA replication and sterol metabolism.

Transcriptomic changes with E disease severity (branch 3)

Differentially expressed genes in disease severity comparison

In all 3 phases (Table 8) in more severe disease, up-regulation of genes for angiogenesis, immune activation, and Fas signaling were observed, with down-regulation of TGF-β signaling and multiple transcription factors. In proliferative phase, more advanced disease was associated with stimulation of genes for fibroblast activation, integrins, collagens, extracellular matrix, MIR21, immune cell activation and cell proliferation (Table 9). Interestingly, several genes with known importance in endometrial function and E were up-regulated in less severe disease, including, MSX2, ILs, BMPs, Indian hedgehog (IHH), and transducer of ERBB signaling (34). In advanced disease genes associated with thyroid hormone metabolism, calcium binding, extracellular matrix, ubiquitin pathway members, and cell cycle regulation were up-regulated in early-secretory phase, whereas in less severe disease, genes for matrix proteins, ion channels, and ephrin receptor were stimulated (Table 9). In advanced disease midsecretory phase genes for phosphoinositol signaling, nitric oxide (NO) synthase trafficking, calcium homeostasis, matrix degradation, angiogenesis, immune modulation, and prostaglandin signaling were up-regulated, and in less severe disease stimulation of genes for IGFBPs, chemokines/receptors, inflammation, and nuclear receptors was observed.

Pathway analyses

In endometrium from women with moderate/severe E, multiple signaling pathways, endothelial biology, inflammation, and cell division were activated in the proliferative and early-secretory phases (Table 7 and Supplemental Table 8), consistent with a previous report (36). In the midsecretory phase, activation of transcription, translation, and pattern binding pathways were observed in moderate/severe disease, whereas cell division, cytokine/growth factor activity, and Wnt signaling were activated in minimal/mild disease, consistent with a previous report (36).

Validation of gene expression

Using multiplex Q-RT-PCR, we validated 92 genes generated from the differentially expressed gene lists corresponding to sample groups at each of the decision tree branches, as well as E vs normal and NE.UPP vs normal. Choice of genes was based on their relevance to core genes in the classifier gene lists, the magnitude of their relative expression, and relevance to endometrial biology (Supplemental Table 11). High concordance was observed between microarray and Q-RT-PCR fold changes, using the Spearman's Rho and Kendall Tau statistics (0.68 and 0.48 for all proliferative phase decision point comparisons, and 0.74 and 0.54 and 0.70 and 0.49 for E vs normal and NE.UPP vs normal, respectively, and all with $P < .0001$).

Discussion

Biomarkers and diagnosing E

Non- or low-invasive biomarkers/diagnostics have the promise of replacing surgical diagnosis of E, although numerous approaches based on candidate genes/proteins relevant to disease pathophysiology and also unbiased approaches (18, 37,–40) have had disappointing results (41). Two recent systematic reviews of all proposed E-related biomarkers (steroid hormone signaling and metabolizing enzymes, cytokines, angiogenic factors, and growth factors) over the last 25 years in serum, plasma, and urine (39) and endometrium (18) could not identify an unequivocally clinically useful biomarker or panel of biomarkers, due to low numbers of subjects, small populations for validations, cycle/hormonal- and stage-dependence, poorly defined controls, and low sensitivity and specificity. A systematic review of more than 200 potential endometrial biomarkers, including hormones and their receptors, cytokines, factors identified through proteomics, and histology revealed sensitivities and specificities of 0%–100% (18). Six high-quality studies identified putative biomarkers related to nerve fiber growth, although recent data suggest that these are a bellwether of pelvic pain and not specifically E (37, 42). Recently, 28 inflammatory and noninflammatory serum biomarkers enabled diagnosis of E in women with disease undetectable by

ultrasound with 81%–90% sensitivity and 63%–80% specificity in independent training and test set data and await further prospective validation (43). Although microRNAs have been discovered in endometrium of women with vs without disease, their utility as a biomarker for E has been suggested, although not specifically tested (44).

Prevalent confounders of uncertain patient characteristics, including menstrual cycle phase, disease stage, comorbid pelvic or uterine conditions, and whether control groups have the absence of uterine and/or other pelvic pathologies are distinct challenges in diagnostic development for E (37). Standard operating procedures for specimen collection, storage, and processing and comprehensive and accurate surgical phenotypic data are vital to the fidelity of the conclusions drawn from the molecular readouts. Herein, all surgeons were experienced in E identification and staging, and we largely followed recent guidelines on standard operating procedures and surgical data collection for E research from a global collaborative (45, 46).

Our classifiers described herein stratified disease categories distinguishing between endometrial tissue samples from women with no abnormalities in the pelvis vs some abnormality (fibroids, E, adenomyosis, pelvic organ prolapse) and also samples from women with E vs other pathology, even when the latter occurred concomitantly with E (a common finding in our cohort). These data suggest that in women with uterine/pelvic abnormalities, the endometrium has a distinct transcriptome compared with samples from women with a completely normal pelvis and that this effect is steroid hormone specific and also independent, underscoring the importance of appropriate stratification in biomarker discovery. Of note, a recent study (47) reported that of 49 endometrial samples analyzed, proteomic, but not transcriptomic, profiles had high accuracy in distinguishing women with E vs no disease. However, the control group included women with and without a normal pelvis, underscoring the importance of defining controls.

Staging of E is surgical, and our classifiers distinguished lesser disease from more advanced disease, not heretofore differentiated with other targeted or unbiased transcriptomic approaches cited above. Knowing this information a priori has the potential to guide treatment options for patients (1). Of interest is a recent urinary peptide profiling study wherein 2 peptide peaks distinguished moderate E from controls and 2 peaks distinguished disease severity (all with 72%–85% sensitivities and specificities) (48). Whether secreted proteins corresponding to genes differentially expressed in our study correspond to as yet unidentified proteomic peaks in serum or urine awaits further investigation.

Molecular and cellular insights derived from unbiased classifier development

Analysis of genes differentially expressed in endometrial samples corresponding to each binary decision in Figure 1B and analysis of their corresponding canonical pathways revealed unanticipated insights into molecular processes and cellular functions in this tissue in the setting of E (and other uterine/pelvic abnormalities), as well as stages of E. Genes uniquely differentially expressed in 1 phase of the cycle provide information about processes that differ between comparison groups in a defined hormonal milieu. However, the discovery of genes that are up- or down-regulated and activation of specific pathways in all 3 phases suggests abnormalities in endometrium that are intrinsic to the presence of uterine and/or pelvic pathologies and are independent of hormonal milieu. In the secretory phase, these changes may also reflect resistance to P_4 action. We observed cycle-independent up-regulation of FOS (5- to 95-fold), a proto-oncogene/transcription factor induced by cytokines, growth factors, and estradiol that modulates expression of genes regulating cell proliferation, differentiation, survival, and angiogenesis (49). FOS is inhibited by P_4 in rat uterus (50) and was highly up-regulated (200-fold) in endometrium in a baboon model of E (51). Why FOS is up-regulated in PE of women with E and other uterine/pelvic pathologies remains to be determined, although the known P_4 resistance in endometrium of women with E (16) may contribute to the apparent cycle-"independent" FOS up-regulation observed herein in the secretory phase.

Also striking in endometrium of women with E was up-regulation across the cycle of genes and pathways involved in estrogen and androgen receptor and growth factor signaling, and immune activation, including antigen presentation, C-C chemokine receptor type 3 (CCR3) eosinophil and CXC chemokine receptor 4 (CXCR4) pathways, lymphocyte activation, cytokine/chemokine activity, humoral immune defense, and the stress/inflammatory response. There was marked up-regulation of IL-32, which induces monocytes and macrophages to secrete proinflammatory cytokines (52), CCL3 (macrophage inflammatory protein1-α), involved in polymorphonuclear leukocyte recruitment/activation in acute inflammation (53), and granulysin, a proinflammatory, cytolytic product of cluster of differentiation 8 (CD8) cytotoxic T cells (54). Whether up-regulation of cytokines, growth factors, and activation of, eg, steroid hormone and/or MAPK signaling pathways, are related to FOS up-regulation awaits further studies, as do identification of the cells of origin of the gene expression signatures and stimuli responsible for the immune activation observed.

Few studies comparing transcriptomes of eutopic endometrium, whole-tissue or isolated cell populations from women with E vs controls have been reported (16,–20, 47), using different platforms and numbers of genes analyzed. Striking, however, is the conserved dysregulation of pathways involved in cell cycle regulation, cell survival, and ERK1/2, MAPK, and PI3K signaling. Similar abnormalities have been found in a baboon model of induced E (35), suggesting that eutopic endometrial changes may be a consequence of the disease. Whether endometrial differential gene expression is a cause or consequence of E (or other uterine/pelvic pathologies), an important biological question, it is not material to the utility of transcriptomic profiling for diagnostic classification. The findings described herein confirm that eutopic endometrium has a consistently unique transcriptomic signature in the setting of E, other gynecologic pathologies, and normal controls.

Clinical implications of endometrial function in women with E and other uterine/pelvic pathologies

Although there are notable differences in the endometrial transcriptome of women with E and those without E but with other uterine/pelvic pathologies, many genes were commonly expressed, supporting the hypothesis that the endometrium is perturbed when some abnormality is present in the pelvis or uterus, including uterine polyps (55), hydrosalpinges (56), and uterine fibroids (57). Whether abnormalities in the uterus and/or pelvis result in or are a result of a proinflammatory environment and perhaps steroid hormone-immune interactions remains to be determined. Furthermore, although the observed perturbations are not identical, the significant overlap supports the need for proper definition of control groups in studies on E biomarkers.

Our observations raise the questions of how abnormal processes (eg, aberrant steroid hormone and growth factor signaling and immune activation), present in nonpregnant endometrium, may affect normal endometrial function, including, eg, embryo implantation and pregnancy outcome and whether some implantation-based pregnancy disorders (eg, preeclampsia, preterm birth, fetal growth restriction) may be mitigated by diagnosis and treatments before pregnancy initiation. The activated signaling pathways may offer opportunities for targeted therapies of endometrial dysfunction before pregnancy or other endometrial disorders unrelated to pregnancy.

With regard to differences in endometrium from women with more advanced vs lesser stage disease, genes and pathways consistent with phase-dependent dysregulation of intracellular signaling pathways, inflammation, and synaptic transmission were observed. These data confirm and extend those of others (18, 36) and suggest an opportunity to discover specific endometrial targets to minimize disease-related pain and infertility.

Limitations

Ultimate clinical usefulness of the discovered classifiers relies on blinded prospective validation in well-defined and larger cohorts. Challenges of further development include potential confounders as endogenous/exogenous hormones and concurrent conditions beyond those stratified in the current analysis. Nonetheless, the current discovery of classifiers able to discriminate endometrial samples from women with E and distinguish stage with high accuracy in the face of existing comorbidities indicates that this approach holds promise for success in development as a clinical diagnostic. A multicenter trial is underway to prospectively validate performance of specific classifiers on an independent test set and determine whether office endometrial biopsy can replace current surgical practice to diagnose and stage disease.

Also, although this study revealed valuable information about underlying biological processes within endometrial tissue in the setting of pelvic disorders, cellular components responsible for these changes remain to be defined by suitable in situ/ex vivo studies. The findings of highly prevalent steroid and thyroid hormone metabolism and signaling and immune activation/inflammation-related pathways provide important clues to cell specificity and narrow the range of potential key cell types involved in these processes.

Summary

In summary, this study lays a foundation for developing low-invasive diagnostics for E and disease staging and also for understanding the roles of the endocrine and immune systems in women with E and other uterine/pelvic disorders and how these affect normal and abnormal endometrial function.

Acknowledgments

We thank the participants who gave informed consent for the use of their tissue in this study, Dr A. Jacoby, Dr M. Goldman, Dr M. Autry, and Dr S. Knight of the University of California for collecting some of the endometrial samples, and the nurses and clinical, laboratory, and administrative staff who made this study possible.

Present address for J.S.T.: Division of Biostatistics, Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA 94305.

This work was supported by the National Institutes of Health, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Specialized Cooperative Centers Program in Reproduction and Infertility Research Grant U54HD 055764 (to L.C.G.), the National Institutes of Health/University of California, San Francisco Human Endometrial Tissue and DNA Bank, the University of California, San Francisco Clinical and Translational Research Institute Resource Allocation Program (L.C.G.), and a University of California Office of the President Proof of Concept grant (L.C.G.).

Disclosure Summary: J.S.T., J.C.I., and L.C.G. are coinventors on provisional United States patent application number 61/885,284. L.C.G. is also an inventor on United States Patent numbers 7,871,778 and 8,247,174 and receives royalties from Wiley Blackwell, Informa, and Elsevier. C.N. receives royalties from McGraw Hill, Springer-Verlag, and Cambridge Press. G.A.G., J.T.R., R.O.B., and L.V.D. have nothing to disclose.

Footnotes

Abbreviations:

BMP bone morphogenetic protein
CCL chemokine (C-C motif) ligand
CYR61 cysteine rich angiogenic inducer 61
cysteine-rich DIO2 deiodinase iodothyronine type II
E endometriosis
E.Min/Mild endometriosis, minimal/mild

E.Mod/Severe endometriosis, moderate/severe
EGF epidermal growth factor
ER estrogen receptor
ERBB v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2
ESE early-secretory phase endometrium
FOSB Fbj Murine Osteosarcoma Oncogene B
GPCR G-protein coupled receptor
GPX glutathione peroxidase
HLA human leukocyte antigen
IGFBP1 insulin-like growth factor binding protein 1
IHH Indian hedgehog
Jak janus kinase
KIR killer immunoglobulin receptor
MSE midsecretory phase endometrium
NCOA2 nuclear receptor coactivator 2
NE.NUPP no E and no uterine/pelvic pathology
NE.UPP no E but with uterine/pelvic pathology
P_4 progesterone
PE proliferative phase endometrium
PI3K phosphatidylinositol 3-kinase
PTEN phosphatase and tensin homologue
q quantitative
STAT signal transducer and activator of transcription
UCSF University of California, San Francisco.

References

1. Giudice LC. Clinical practice. Endometriosis. N Engl J Med. 2010;362(25):2389–2398. [PMCID: PMC3108065] [PubMed: 20573927]

2. Giudice LC, Kao LC. Endometriosis. Lancet. 2004;364(9447):1789–1799. [PubMed: 15541453]

3. Sampson JA. Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity. Am J Obstet Gynecol. 1927;14:442–469.

4. Berkley KJ, Rapkin AJ, Papka RE. The pains of endometriosis. Science. 2005;308(5728):1587–1589. [PubMed: 15947176]

5. Tokushige N, Markham R, Russell P, Fraser IS. Nerve fibres in peritoneal endometriosis. Hum Reprod. 2006;21(11):3001–3007. [PubMed: 16950827]

6. Eskenazi B, Warner ML. Epidemiology of endometriosis. Obstet Gynecol Clin North Am. 1997;24(2):235–258. [PubMed: 9163765]

7. Meuleman C, Vandenabeele B, Fieuws S, Spiessens C, Timmerman D, D'Hooghe T. High prevalence of endometriosis in infertile women with normal ovulation and normospermic partners. Fertil Steril. 2009;92(1):68–74. [PubMed: 18684448]

8. Nnoaham KE, Hummelshoj L, Webster P, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. Fertil Steril. 2011;96(2):366–373.e368. [PMCID: PMC3679489] [PubMed: 21718982]

9. Simoens S, Dunselman G, Dirksen C, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Hum Reprod. 2012;27(5):1292–1299. [PubMed: 22422778]

10. Bulun SE. Endometriosis. N Engl J Med. 2009;360(3):268–279. [PubMed: 19144942]

11. Hadfield R, Mardon H, Barlow D, Kennedy S. Delay in the diagnosis of endometriosis: a survey of women from the USA and the UK. Hum Reprod. 1996;11(4):878–880. [PubMed: 8671344]

12. D'Hooghe TM, Debrock S. Endometriosis, retrograde menstruation and peritoneal inflammation in women and in baboons. Hum Reprod Update. 2002;8(1):84–88. [PubMed: 11866244]

13. Ballard K, Lowton K, Wright J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. Fertil Steril. 2006;86(5):1296–1301. [PubMed: 17070183]

14. Melin AS, Lundholm C, Malki N, Swahn ML, Sparèn P, Bergqvist A. Hormonal and surgical treatments for endometriosis and risk of epithelial ovarian cancer. Acta Obstet Gynecol Scand. 2013;92(5):546–554. [PubMed: 23560387]

15. D'Hooghe TM, Mihalyi AM, Simsa P, et al. Why we need a noninvasive diagnostic test for minimal to mild endometriosis with a high sensitivity. Gynecol Obstet Invest. 2006;62(3):136–138. [PubMed: 16679771]

16. Burney RO, Talbi S, Hamilton AE, et al. Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis. Endocrinology. 2007;148(8):3814–3826. [PubMed: 17510236]

17. Kao LC, Germeyer A, Tulac S, et al. Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility. Endocrinology. 2003;144(7):2870–2881. [PubMed: 12810542]

18. May KE, Villar J, Kirtley S, Kennedy SH, Becker CM. Endometrial alterations in endometriosis: a systematic review of putative biomarkers. Hum Reprod Update. 2011;17(5):637–653. [PubMed: 21672902]

19. Matsuzaki S, Canis M, Vaurs-Barriere C, Boespflug-Tanguy O, Dastugue B, Mage G. DNA micro array analysis of gene expression in eutopic endometrium from patients with deep endometriosis using laser capture microdissection. Fertil Steril. 2005;84(suppl 2):1180–1190. [PubMed: 16210010]

20. Sherwin JR, Sharkey AM, Mihalyi A, Simsa P, Catalano RD, D'Hooghe TM. Global gene analysis of late secretory phase, eutopic endometrium does not provide the basis for a minimally invasive test of endometriosis. Hum Reprod. 2008;23:1063–1068. [PubMed: 18353903]

21. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. Fertil Steril. 1997;67(5):817–821. [PubMed: 9130884]

22. Noyes RW, Hertig AT, Rock J. Dating the endometrial biopsy. Am J Obstet Gynecol. 1975;122(2):262–263. [PubMed: 1155504]

23. Talbi S, Hamilton AE, Vo KC, et al. Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. Endocrinology. 2006;147(3):1097–1121. [PubMed: 16306079]

24. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F. A model-based background adjustment for oligonucleotide expression arrays. J Am Stat Assoc. 2004;99(468):909–917.

25. Tibshirani R, Hastie T. Margin trees for high-dimensional classification. J Mach Learn Res. 2007;8:637–652.

26. Pepe MS, Longton G, Anderson GL, Schummer M. Selecting differentially expressed genes from microarray experiments. Biometrics. 2003;59(1):133–142. [PubMed: 12762450]

27. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3:Article3.

28. Tian L, Greenberg SA, Kong SW, Altschuler J, Kohane IS, Park PJ. Discovering statistically significant pathways in expression profiling studies. Proc Natl Acad Sci USA. 2005;102(38):13544–

13549. [PMCID: PMC1200092] [PubMed: 16174746]

29. Spitzer TL, Rojas A, Zelenko Z, et al. Perivascular human endometrial mesenchymal stem cells express pathways relevant to self-renewal, lineage specification, and functional phenotype. Biol Reprod. 2012;86(2):58. [PMCID: PMC3290674] [PubMed: 22075475]

30. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 2008;3(6):1101–1108. [PubMed: 18546601]

31. Rorabacher DB. Statistical treatment for rejection of deviant values: critical values of Dixon's “Q” parameter and related subrange ratios at the 95% confidence level. Anal Chem. 1991;63(2):139–146.

32. Lee S. Mistakes in validating the accuracy of a prediction classifier in high-dimensional but small-sample microarray data. Stat Methods Med Res. 2008;17(6):635–642. [PubMed: 18375459]

33. Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. J Clin Oncol. 2005;23(29):7332–7341. [PubMed: 16145063]

34. Altmäe S, Esteban FJ, Stavreus-Evers A, et al. Guidelines for the design, analysis and interpretation of ‘omics’ data: focus on human endometrium. Hum Reprod Update. 2014;20(1):12–28. [PMCID: PMC3845681] [PubMed: 24082038]

35. Afshar Y, Hastings J, Roqueiro D, Jeong JW, Giudice LC, Fazleabas Changes in eutopic endometrial gene expression during the progression of experimental endometriosis in the baboon, *Papio anubis*. Biol Reprod. 2013;88(2):44–53. [PMCID: PMC3589234] [PubMed: 23284138]

36. Aghajanova L, Giudice LC. Molecular evidence for differences in endometrium in severe versus mild endometriosis. Reprod Sci. 2011;18(3):229–251. [PMCID: PMC3118406] [PubMed: 21063030]

37. Fassbender A, Vodolazkaia A, Saunders P, et al. Biomarkers of endometriosis. Fertil Steril. 2013;99(4):1135–1145. [PubMed: 23414923]

38. Fassbender A, Waelkens E, Verbeeck N, et al. Proteomics analysis of plasma for early diagnosis of endometriosis. Obstet Gynecol. 2012;119(2 pt 1):276–285. [PubMed: 22270279]

39. May KE, Conduit-Hulbert SA, Villar J, Kirtley S, Kennedy SH, Becker CM. Peripheral biomarkers of endometriosis: a systematic review. Hum Reprod Update. 2010;16(6):651–674. [PMCID: PMC2953938] [PubMed: 20462942]

40. Zheng N, Pan C, Liu W. New serum biomarkers for detection of endometriosis using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Int Med Res. 2011;39(4):1184–1192. [PubMed: 21986120]

41. Rogers PA, D'Hooghe TM, Fazleabas A, et al. Defining future directions for endometriosis research: workshop report from the 2011 World Congress of Endometriosis In Montpellier, France. Reprod Sci. 2013;20(5):483–499. [PMCID: PMC3635070] [PubMed: 23427182]

42. Zhang X, Lu B, Huang X, Xu H, Zhou C, Lin J. Innervation of endometrium and myometrium in women with painful adenomyosis and uterine fibroids. Fertil Steril. 2010;94(2):730–737. [PubMed: 19368910]

43. Vodolazkaia A, El-Aalamat Y, Popovic D, et al. Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. Hum Reprod. 2012;27(9):2698–2711. [PubMed: 22736326]

44. Pan Q, Luo X, Toloubeydokhti T, Chegini N. The expression profile of micro-RNA in endometrium and endometriosis and the influence of ovarian steroids on their expression. Mol Hum Reprod. 2007;13(11):797–806. [PubMed: 17766684]

45. Fassbender A, Rahmioglu N, Vitonis AF, et al. for the WERF EPHect Working Group. WERF Endometriosis Phenome and Biobanking Harmonization Project (EPHect): IV. Tissue collection, processing and storage in endometriosis research [published online September 22, 2014]. Fertil Steril. doi:10.1016/j.fertnstert.2014.07.1209. [PMCID: PMC4230778]

46. Becker CM, Laufer MR, Stratton P, et al. for the WERF EPHect Working Group. WERF Endometriosis Phenome and Biobanking Harmonization Project (EPHect): I. Surgical phenotype data collection in endometriosis research [published online August 21, 2014]. Fertil Steril. doi:10.1016/j.fertnstert.2014.07.709. [PMCID: PMC4230690]

47. Fassbender A, Verbeeck N, Börnigen D, et al. Combined mRNA microarray and proteomic analysis of eutopic endometrium of women with and without endometriosis. Hum Reprod. 2012;27(7):2020–2029. [PubMed: 22556377]

48. El-Kasti MM, Wright C, Fye HK, Roseman F, Kessler BM, Becker CM. Urinary peptide profiling identifies a panel of putative biomarkers for diagnosing and staging endometriosis. Fertil Steril. 2011;95(4):1261–1266.e1–e6. [PubMed: 21255776]

49. Hu E, Mueller E, Oliviero S, Papaioannou VE, Johnson R, Spiegelman BM. Targeted disruption of the c-fos gene demonstrates c-fos-dependent and -independent pathways for gene expression stimulated by growth factors or oncogenes. EMBO J. 1994;13(13):3094–3103. [PMCID: PMC395200] [PubMed: 8039503]

50. Kirkland JL, Murthy L, Stancel GM. Progesterone inhibits the estrogen-induced expression of c-fos messenger ribonucleic acid in the uterus. Endocrinology. 1992;130(6):3223–3230. [PubMed: 1375896]

51. Hastings JM, Jackson KS, Mavrogianis PA, Fazleabas AT. The estrogen early response gene FOS is altered in a baboon model of endometriosis. Biol Reprod. 2006;75(2):176–182. [PubMed: 16672717]

52. Kim SH, Han SY, Azam T, Yoon DY, Dinarello CA. Interleukin-32: a cytokine and inducer of TNFα. Immunity. 2005;22(1):131–142. [PubMed: 15664165]

53. Menten P, Wuyts A, Van Damme J. Macrophage inflammatory protein-1. Cytokine Growth Factor Rev. 2002;13(6):455–481. [PubMed: 12401480]

54. Janeway C. Immunobiology: The Immune System in Health and Disease. 6th ed New York, NY: Garland Science; 2005.

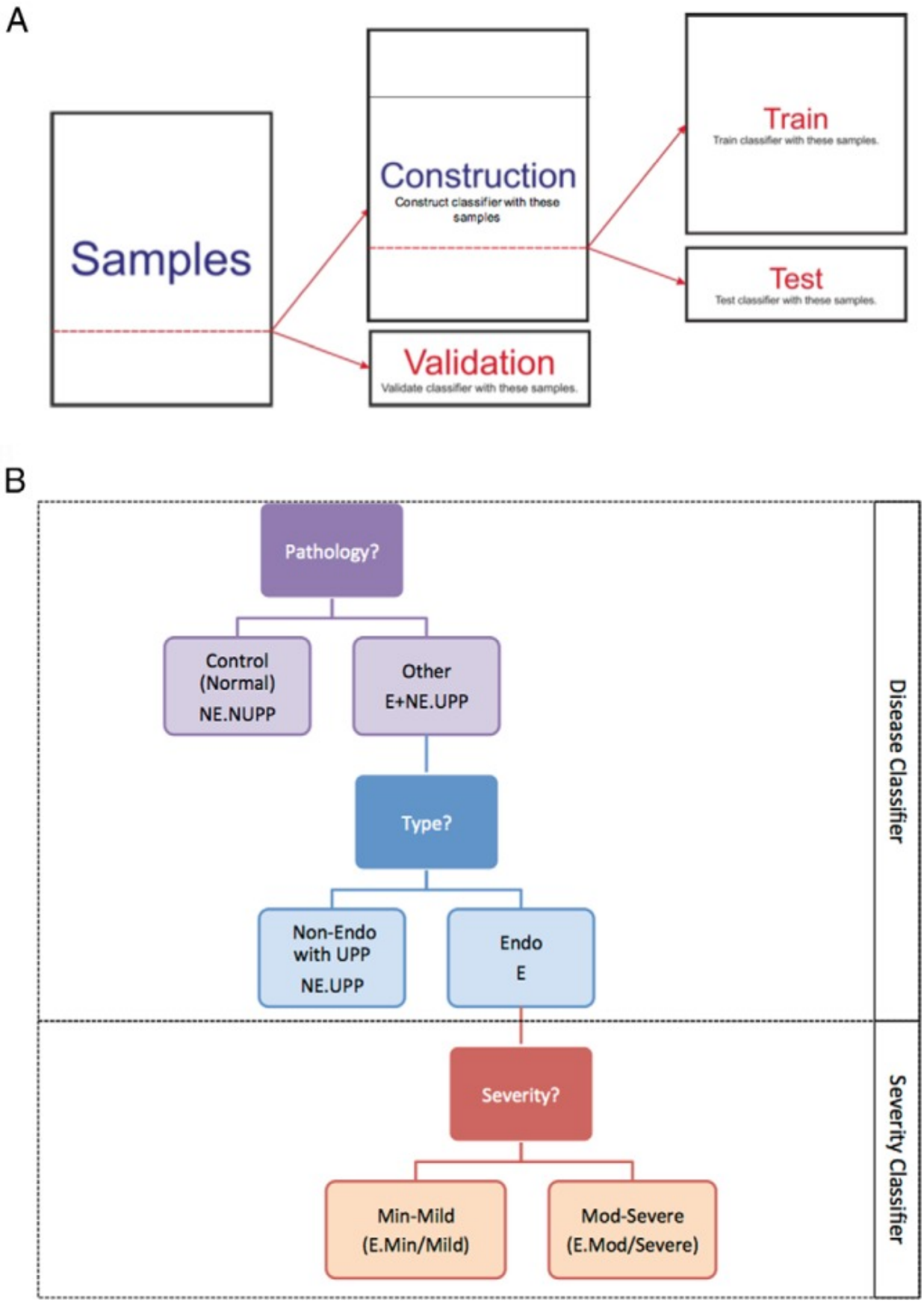
55. Rackow BW, Jorgensen E, Taylor HS. Endometrial polyps affect uterine receptivity. Fertil Steril. 2011;95(8):2690–2692. [PMCID: PMC3096716] [PubMed: 21269620]

56. Meyer WR, Castelbaum AJ, Somkuti S, et al. Hydrosalpinges adversely affect markers of endometrial receptivity. Hum Reprod. 1997;12(7):1393–1398. [PubMed: 9262264]

57. Rackow BW, Taylor HS. Submucosal uterine leiomyomas have a global effect on molecular determinants of endometrial receptivity. Fertil Steril. 2010;93(6):2027–2034. [PMCID: PMC3107853] [PubMed: 18555231]

Figures and Tables

Figure 1.


A
Samples
Construction
Construct classifier with these samples
Train
Train classifier with these samples.
Test
Test classifier with these samples.
Validation
Validate classifier with these samples.
B
Pathology?
Control (Normal)
NE.NUPP
Other
E+NE.UPP
Type?
Non-Endo with UPP
NE.UPP
Endo
E
Severity?
Min-Mild
(E.Min/Mild)
Mod-Severe
(E.Mod/Severe)
Disease Classifier
Severity Classifier

Classifier development. A, Schematic representation of classifier construction and validation showing partitioning of samples to construct classifiers, including train and test sets, and validating classifiers with samples not used in the construction process. B, Binary decision tree to diagnose and stage E. min-mild, minimal/mild; mod-severe, moderate/severe.

Table 1.

Endometrial Samples by Menstrual Cycle Phase, Group, and Disease Severity

Group		PE	ESE	MSE	Total
NE.NUPP		20	6	8	34
NE.UPP		15	6	14	35[a]
E		29	18	28	75[b]
	Minimal/mild (E.Min/Mild)	11	6	10	27[b]
	Moderate/severe (E.Mod/Severe)	18	12	18	48

[a]Two additional NE.UPP samples had indeterminate cycle phase assignment.
[b]Two additional E samples had indeterminate cycle phase assignment: 1 E.Min/Mild and 1 E.Mod/Severe sample had indeterminate cycle phase assignments.

Table 2.

Classifier Performance. Performance of Disease Classifiers

Diagnostic Variant	Cycle Phase	Construction Samples	Validation Samples	Cross-Validation Folds	Classifiers/Family	Validation Accuracy
Phase unrestricted	PE + ESE + MSE	120	28	10	4	93%
Phase restricted	PE + ESE	76	18	10	2	100%
Phase specific	PE	51	13	10	11	100%
	ESE	24	6	5	54	100%
	MSE	39	11	6	4	91%

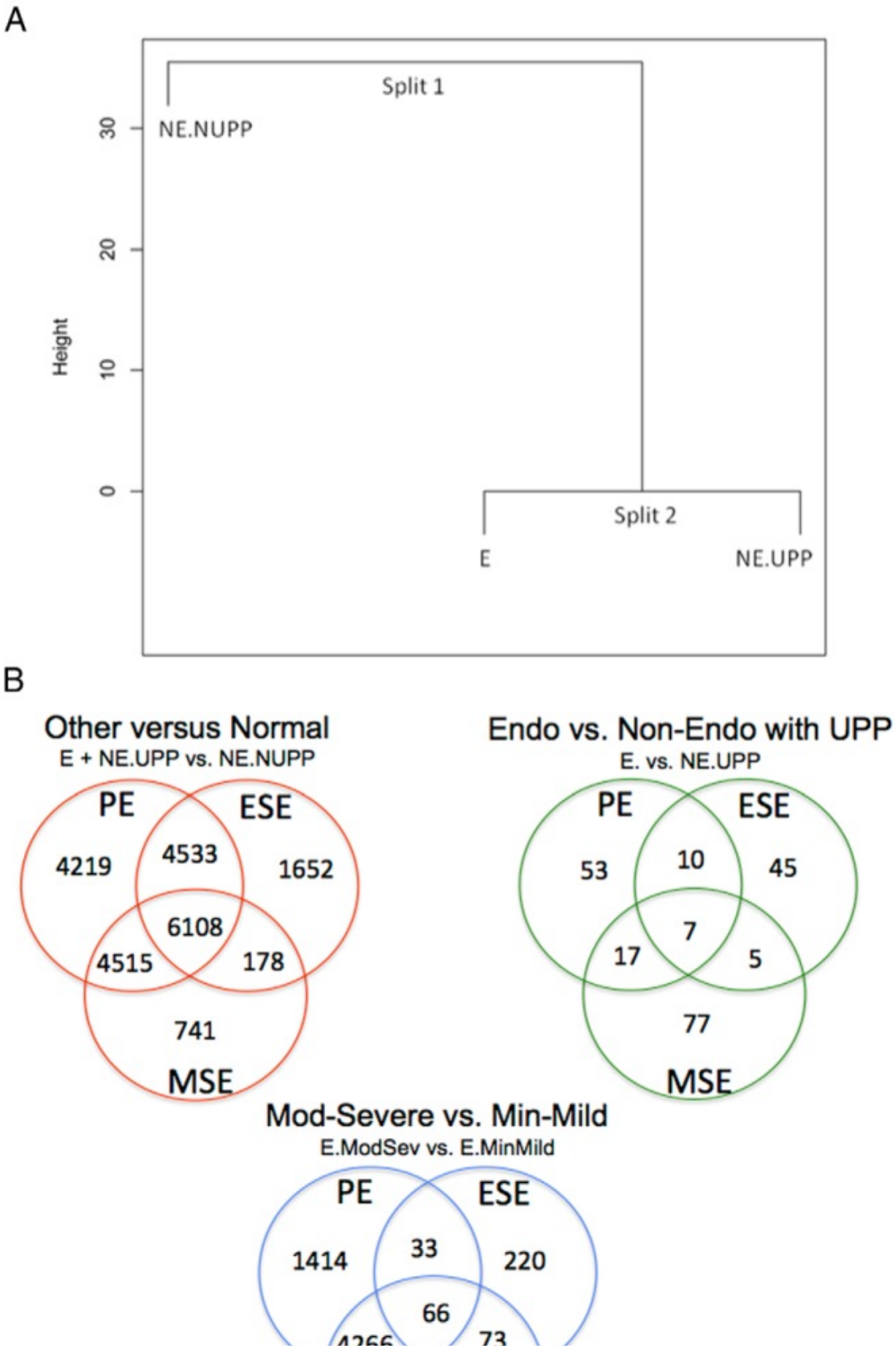
Phase unrestricted analysis includes samples from all phases: PE, ESE, and MSE. Phase restricted analysis includes PE and ESE but does not include MSE. Phase specific analysis includes PE or ESE or MSE.

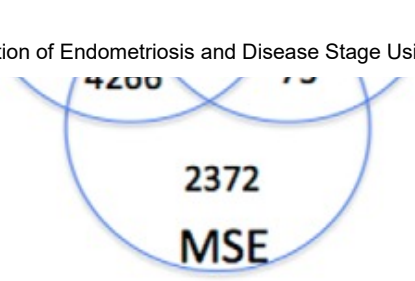
Table 3.

Classifier Performance. Performance of Phase-Specific Severity Classifiers

Diagnostic Variant	Cycle Phase	Construction Samples	Validation Samples	Cross-Validation Folds	Classifiers/Family	Validation Accuracy
Phase specific	PE	22	6	9	43	100%
	ESE	14	4	5	22	100%
	MSE	21	6	5	44	100%

Figure 2.


A
Split 1
NE.NUPP
Height
30
20
10
0
Split 2
E
NE.UPP
B
Other versus Normal
E + NE.UPP vs. NE.NUPP
PE
ESE
MSE
4219
4533
1652
6108
4515
178
741
Endo vs. Non-Endo with UPP
E. vs. NE.UPP
PE
ESE
MSE
53
10
45
7
17
5
77
Mod-Severe vs. Min-Mild
E.ModSev vs. E.MinMild
PE
ESE
1414
33
220
66
73



A, E disease classifier margin tree dendrogram. B, Venn diagrams showing number of differentially expressed probe sets in the comparison groups shown. E.ModSevere, moderate/severe stage endometriosis; E.MinMild, minimal/mild stage endometriosis.

Table 4.

Select Genes Differentially Expressed (fold change) Across All Cycle Phases: Abnormal (E and/or NE.UPP) vs Normal (NE.NUPP)

Phase	Gene	All Abnormal (E + NE.UPP) vs Normal (NE.NUPP)			E vs Normal (NE.NUPP)			Non-E (NE.UPP) vs Normal (NE.NUPP)		
		PE	ESE	MSE	PE	ESE	MSE	PE	ESE	MSE
Up										
	FOSB	48.6	12.4	17.6	34.1	10.2	17.0	96.5	22.1	17.1
	FOS	33.5	30.0	38.0	25.8	29.1	33.7	55.6	32.9	48.1
	EGR1	16.2	30.5	12.7	15.2	31.4	13.6	25.5	12.8	17.8
	JUNB	15.3	6.3	5.0	13.1	6.0	5.5	20.4	7.4	4.2
		14.7	6.1	4.5	11.4	5.2	3.6	24.2	9.5	6.9
MTSS1L										
	CTSW	10.0	5.4	6.9	7.6	5.6	6.6	16.9	4.9	7.4
		9.4	5.4	4.3	7.8	5.1	4.7	13.4	6.2	7.0
TGFB1										
	SOC3	9.3	3.5	3.1	7.6	3.0	3.2	13.9	5.7	2.7
	IL32	7.8	3.8	6.1	6.0	3.8	5.3	13.7	3.9	8.1
		7.7	4.7	2.7	6.9	4.2	2.5	9.6	6.4	3.1
FKBP8										
		7.3	2.7	3.3	6.4	2.4	2.9	9.3	3.8	4.4
ISYNA1										
	CCL3	7.2	4.1	7.5	5.2	3.9	6.5	13.2	4.9	10.0
	GNLY	6.0	5.4	2.6	4.8	6.1	2.7	6.7	2.8	2.5
		6.0	3.0	3.5	5.6	2.6	3.1	6.6	4.6	4.2
MAP3K11										
	C1QA	5.6	2.5	7.3	5.2	2.5	5.9	6.6	2.6	11.0
		4.4	2.7	3.0	3.8	1.9	1.9	6.1	3.9	3.7
NOTCH3										
		3.6	3.3	7.9	1.9	3.0	6.3	12.1	4.3	12.5
CYR61										
		3.5	2.9	2.8	3.1	2.7	2.6	4.3	3.3	5.7
NPTXR										
Down										
	FBN1	−8.7	−2.8	−3.6	−9.0	−2.4	−3.4	−8.1	−4.6	−4.0
		−8.7	−3.6	−5.2	−7.2	−3.0	−3.0	−12.5	−6.2	−8.7
PNRC2										
		−8.8	−5.1	−3.1	−8.0	−4.4	−2.9	−10.8	−7.9	−3.7
ITGA6										
	DHFR	−8.8	−5.1	−2.1	−9.0	−5.0	−2.1	−8.3	−5.3	−2.1

Phase Gene	All Abnormal (E + NE.UPP) vs Normal (NE.NUPP)			E vs Normal (NE.NUPP)			Non-E (NE.UPP) vs Normal (NE.NUPP)		
	PE	ESE	MSE	PE	ESE	MSE	PE	ESE	MSE
SLC39A6	−9.1	−4.5	−3.5	−8.1	−3.7	−3.2	−11.6	−8.3	−3.8
MYO10	−9.5	−4.5	−5.6	−9.2	−3.8	−2.9	−9.8	−4.3	−4.9
HSP90B1	−10.6	−4.1	−2.9	−8.1	−3.7	−2.9	−17.8	−5.9	−3.0
SMC3	−11.2	−5.2	−3.3	−11.7	−4.1	−3.6	−10.3	−10.9	−2.8
PKP4	−12.0	−5.0	−4.8	−10.6	−4.7	−4.2	−15.3	−5.8	−6.1
PALLD	−12.7	−4.6	−2.6	−14.3	−4.6	−2.6	−10.1	−4.4	−2.5
DIO2	−15.6	−3.4	−5.1	−19.3	−4.6	−5.7	−10.3	−8.5	−4.1

EGR1, early growth response 1.

Table 5.

Select Genes Uniquely Differentially Expressed (fold change) in Each Cycle Phase: Abnormal ([E + NE.UPP]) vs Normal (NE.NUPP)

	PE		ESE		MSE			
Up	GPX3	10.3	NAMPT	6.5	PLAG2A	6.1	KIR2DL3	3.0
	AQP3	6.5	PDZD2	4.6	IGFBP5	6.1	HBA1/A2	2.8
	GDF15	5.8	TMEM49	4.4	KIR3DL1/2	4.5	ANG	2.6
	LCN2	5.8	COL6A1	4.3	KIR2DS2	4.0	IL10RA	2.5
	FOSL2	5.3	PTEN	4.0	THY1	4.0	KIR2DS1	2.5
	FOLR1	5.0	PSME4	3.7	KIR2DL4/5A	3.8	IL6	2.3
	CLIC3	5.0	PPA2	3.6	KIR3DL3	3.6	KIR2DS1	2.2
	SCGB3A1	4.7	PCSK6	3.6	HIST1H4C	3.5	CD44	2.0
	HIST2H2A3/4	4.7	CBFA2T2	3.5	RGS16	3.3	GPX2	2.0
	KRT 17	4.6	ODS5B	3.5	KIR2DL1	3.2	PLAUR	2.0
	MUC5B	4.7	MOGAT1	3.4	CX3CR1	3.0		
	CEBPD	4.6	ZNF638	3.4				
	HLA-DQB1	4.0						
	WNT4	3.8						
Down	WIF1	−4.5	HDAC9	−2.2	NCOA2	−1.9	MUC15	−3.6
	F2RL2	−5.2	C6	−2.2	IL1R1	−2.6	PTGS	−3.6
	KMO	−6.1	EPHX2	−2.2	PRLR	−2.8	TLR5	−3.8
	ANK2	−6.1	EPHA	−2.3	FKBP5	−2.8	CRISP2	−4.3
	CPM	−6.2	PDE3B	−2.4	ADAMTS5	−2.9	ENPP1	−4.3
			THRB6	−2.4	IL6ST	−2.9	QSER	−4.5
			SLC28A3	−2.5	ACPP	−3.0	MYOCD1	−5.0
			OLFM1	−2.7	CDH1	−3.2	SYNE2	−5.4
			PACRG	−2.9	ADAMTS1	−3.3	IDO	−6.5
			SLC26A4	−3.0	GPR64	−3.3	RIMKLB	−9.1
					PAX8	−3.6		

Table 6.

Select Genes Uniquely Differentially Expressed (fold change) in Each Cycle Phase: E vs no E but with Uterine/Pelvic Abnormalities (NE.UPP)

	PE		ESE		MSE	
Up	RPS17	1.6	TMED4	3.3	SLC1A1	5.6
	PQBP1	1.5	MRPS18	2.3	CYP3A5	4.9
			TMEM50A	1.9	EIF1	3.0
			SDF2	1.6	PAX8	3.0
					ABCC3	2.9
					GPR110	2.6
					VEGF8	2.5
					CAPN	1.6
					CLDN10	1.5
Down	CYR61	−6.3	ZNF87	−3.2	DST	−3.9
	MGP	−4.8	USP36	−3.0	RNF150	−3.8
	END1	−3.5	PTK2	−2.8	PCSK5	−3.6
	ACTA2	−3.0	SR140	−2.5	THRAP3	−2.5
	OLML1	−2.2	NUB1	−2.2	MUC7	−2.5
	SLC1A6	−2.0	SCN11A	−1.9	CHD4	−2.4
	DKK3	−1.8	NTN5	−1.8	LDLR	−2.3
	DNMT3A	−1.7			OSBPL8	−2.2
	NAV1	−1.7			SLC8A1	−2.2
	NAV3	−1.7			MLL	−1.5
					EPASL	−1.5

Table 7.

Activated and Inhibited Pathways for Differentially Expressed Genes

Phase/Comparison		E + NE.UPP vs NE.NUPP	E vs NE.NUPP	NE.UPP vs NE.NUPP	E vs NE.UPP	E.Mod.Sev vs E.MinMild
PE						
	Activated	Neuropeptide signaling, immune response, antigen presentation, steroid metabolism, MAPK signaling	Immune activation (lymphocytes, antigen presentation, cytokine activity)	Oxidative stress response, neuronal differentiation, ER, cAMP/Ca^{2+} signaling, immune activation, hypoxia, angiogenesis	None	Jak-STAT, EGF/PDGF, PI3K-AKT signaling, endothelial cell biology, inflammation, protein synthesis, cell division
	Inhibited	Cell division	Cell division, protein synthesis, IFNg and ras signaling, histone binding	Cell division, protein synthesis	Integrin/cell adhesion, chemotaxis, response to chemical substance, stress/toxicity, insulin, hypoxia, and cAMP/Ca^{2+} signaling, IFNab response	Synaptic transmission, cytokines, cation channel activity
ESE						
	Activated	EGF/PDGR, GPCR, ER, AR, MAPK signaling, angiogenesis	Erk1/Erk2, MAPK, chemokine, Wnt, hypoxia, GPCR, AR, ER signaling, immune activation, angiogenesis, integrin/cell adhesion	GPCR, cAMP/Ca^{2+}, Erk1/Erk2, PTEN, chemokine, MAPK, TGFb/BMP, NFkB signaling, immune activation, integrin/cell adhesion	Oxidative phosphorylation, RNA slicing, transporter activity	p53 signaling, integrin-mediated cell adhesion
	Inhibited	Active transport, oxidative phosphorylation, cell division, protein synthesis	Genomic stability/DNA repair, electron transport, growth	Genomic stability/DNA repair, electron transport, cell cycle	None	Muscle development

Phase/Comparison		E + NE.UPP vs NE.NUPP	E vs NE.NUPP	NE.UPP vs NE.NUPP	E vs NE.UPP	E.Mod.Sev vs E.MinMild
MSE						
	Activated	Cell defense, immune cell activation	NO, immune cell activation, cytokine activity, humoral defense, immune activation, stress/inflammatory response, GPCR, MAPK, NFkB, Wnt, hypoxia, growth factor signaling	Oxidative phosphorylation, reactive oxygen species metabolism, response to stress/toxicity, Wnt, MAPK signaling	Response to chemical substance, ER and Wnt signaling, actin cytoskeleton organization and biosynthesis, angiogenesis	Gene transcription regulation, translation, cation channels, pattern binding
	Inhibited	Steroid metabolism, protein kinase activity, negative cell cycle regulation	JakSTAT, thyroid hormone, AR signaling, transcription, steroid metabolism, apoptosis	Gene transcription regulation, translation, Jak-STAT, EGF/PDGF signaling, TLR, secretory pathway	DNA replication, cholesterol and alcohol metabolism	Cell division, phosphate transport, cytokine, growth factor activity, Wnt signaling

Table 8.

Select Genes Differentially Expressed (fold change) Across All Cycle Phases: Moderate/Severe vs Minimal Mild E

Gene	PE	ESE	MSE	Gene	PE	ESE	MSE	Gene	PE	ESE	MSE
EDNRA	3.2	4.4	3.2	ZNF140	1.7	1.8	1.7	TABLN	−2.4	−2.8	−2.3
NFIB	2.6	4.7	3.6	SNAP23	1.7	1.8	2.5	ZNF827	−2.4	−2.9	−5.1
DDX3X	2.5	2.8	3.7	CAPN3	1.6	2.3	1.7	KCNE3	−2.3	−2.8	−2.2
PIK3C2A	2.1	2.6	2.8	G3BP2	1.6	1.8	1.9	MYCL1	−2.0	−2.0	−1.6
RAB5A	2.1	2.5	3.1	TBRG1	1.5	1.7	1.9	SP3	−1.9	−2.5	−2.7
FAS	1.8	2.0	2.1	UQCC	1.5	1.6	1.6	RBP6	−1.9	−1.8	−1.7
RORA	1.8	2.0	1.7	RAP3	1.5	1.6	1.6	CPPED1	−1.8	−2.0	−2.3
MYO5B	1.8	2.2	2.5	TGIF2	−3.4	−1.7	−2.2	ERGIC1	−1.8	−2.3	−2.7
Gene	**PE**	**ESE**	**MSE**	**Gene**	**PE**	**ESE**	**MSE**	**Gene**	**PE**	**ESE**	**MSE**
PRKAA1	1.8	1.8	1.9	MSLN	−3.3	−1.7	−2.2	INSIG1	−1.6	−1.6	−2.1
PTPRM	1.8	2.0	2.0	IFI6	−2.9	−5.0	−6.6				

Table 9.

Select Genes Uniquely Differentially Expressed (fold change) in Each Cycle Phase: Moderate/Severe vs Minimal/Mild E

	Proliferative				Early Secretory		Midsecretory			
Up	FAP	5.6	ZNF117	2.4	DIO2	9.7	RIMKL	12.5	CDH1	4.0
	ITGB1	5.0	BMPR1A	2.4	CALB2	7.8	SLC1A1	11.4	CYP3A5	4.0
	COL12A1	3.0	GABABR1	2.4	COL1A1	4.9	PIK3R1	9.1	FGF7/KGFLP1/2	3.7
	ADAMTS1	4.0	IL13RA1	2.4	TMEM49	4.7	RBM6	9.8	ANGPTL1	3.6
	MME	3.5	FN1	2.3	USP9X	2.9	DPP4	6.9	IDO1	3.2
	CDH11	3.3	EPS8	2.3	CLIC4	2.7	MYOCD	6.6	KDM5	2.9
	ITGA4	3.0	CD59	2.3	FN1	1.9	GBP1	6.0	VCAN	2.8
	DST	2.9	MYST4	2.5	SGRAP1	1.9	CP	5.9	KLF4	2.8
	FOXN3	2.9	BCAP29	2.2	CDK11A/B	1.6	SLC18A2	5.5	IL1R1	2.8
	MIR21	2.9	ITGB8	2.2	GREB1L	1.6	NOSTRIN	5.4	LACTB2	2.8
	VEZF1	2.8	IL10RB	2.1	IL1A	1.6	CALD1	5.2	PIBF1	2.7
	SMC4	2.8	IGF1R	2.0			NFKBIZ	5.1	HSD17B4	2.5
	PSME4	2.8					ADAMTS1	5.0	TNCOA2	2.3
	PDGFRA	2.8					MAOA	4.9	INSR	2.3
	SPOCK1	2.7					C4BPA	4.8	HMGB1	2.3
	MYO10	2.7					IL6ST	4.6	GA6	2.2
	MAL2	2.6					1L15	4.2	PTGER3	2.2
	SPARC	2.4					MAP3K5	4.0	LCP1	2.2
Down	GTSE1	−7.0	GRK6	−2.4	CRISP3	−6.8	IGFBP5	−14.0	CCND3	−3.5
	MYCN	−6.4	NFATC2IP	−2.4	TJSD4	−4.7	MFAP2	−10.4	MUC1	−3.5
	CLIC3	−4.8	GZMM	−2.4	ARJGAP32	−4.0	LOXL1	−9.1	DNMT3A	−3.5
	MSX2	−4.6	FGFR4	−2.3	CTFR	−3.8	RNF24	−7.4	NR1H2	−3.4
	IL32	−4.6	RGS12	−2.3	EFNA5	−3.6	PMEPA1	−6.2	HLA − DMA	−3.4
	GPER	−4.6	IL11RA	−2.3	NRXN3	−3.5	CPZ/GPR78	−5.7	THY1	−3.2
	BMP7	−4.4	MT3	−2.3	CD55	−3.2	NPTX2	−5.6	CXCR7	−3.2
	GFGR3	−4.1	F12	−2.1	RORA	−2.2	IGFBP6	−5.2	BAD	−3.1
	HISTH2A	−4.1	NCS1	−2.1	KDM4C	−2.0	CYGB	−4.8	NFKBIE	−3.1
	FOSL	−3.4	IFNAR1	−2.0			LTBP2	−4.8	ISGF8	−3.0
	IHH	−3.3					DST	−4.8	ITGB4	−3.0
	TOB2	−3.1					BMP2	−4.3	PCGF2	−3.0
	HSD17B8	−3.0					CCL5	−4.3	HEBP1	−3.0
	IFI27L2	−2.8					ICAM2	−4.4	CDKN1C	−2.8
	HSD17B14	−2.7					VWF	−3.8	COL6A2	−2.8
	LCK	−2.5					SDFL1	−3.6	NR1DI/THRA	−2.6
	KLF16	−2.5					ESAM	−3.6	ECSCR	−2.5

Proliferative		Early Secretory	Midsecretory		
FKBP11	−2.5		HGF	−3.5	TIMP1	−2.4