ABSTRACT

Endometriosis is characterized by endometrial-like tissue outside the uterine cavity, derived largely from abnormal eutopic endometrium (within the uterus) refluxed during menses that implants on pelvic organs because of its enhanced survival, angiogenic, and proliferative potential, and eliciting an inflammatory response and attendant infertility and chronic pelvic pain [1]. Eutopic endometrium of women with disease has markedly different transcriptomic and proteomic profiles, abnormal steroid hormone responses (to estradiol [E2] and progesterone [P4] resistance), aberrant growth factor signaling, and a proinflammatory phenotype, compared with unaffected women [2–10]. Although the pathogenesis of these differences is not well understood, epigenetic abnormalities have been implicated [3].

Epigenetics refers to modifications of gene activity that are not accompanied by changes in gene sequence. Methylation of the 5′ carbon position of cytosines, usually in the context of CpG dinucleotides, is the main epigenetic modification of DNA with essential roles in various biological processes [11]. Up to 20% of genes display DNA methylation patterns in a tissue-specific manner, associated with tissue-specific gene expression [12]. The distribution of DNA methylation across the genome and the position of methylation in the transcriptional unit should be considered when assessing DNA methylation data. In the vertebrate genome, more than half of the genes contain CpG islands (CGIs)—short, CG-rich regions—although the rest of the genome is CpG depleted [13]. At the promoters, the CpG density has a bimodal distribution, although intermediate CpG densities also exist [13]. The position of the methylation with relation to the transcriptional unit impacts its relationship with regulation of gene expression [13]. Most CGIs are unmethylated in somatic cells, with only a minority being methylated. The CGIs near the transcription start sites of active genes are usually unmethylated, and de novo methylation at these CGIs, as shown in cancer and disease, is often associated with gene silencing. We have found in women without endometriosis that endometrial DNA methylation changes across the menstrual cycle, involving both CGIs and non-CGIs, and is associated with changes in gene expression for several loci [14]. Greatest differences were observed between the proliferative (peak E2) and the midsecretory (peak P4) phases, suggesting that epigenetic modifications play a role in normal endometrial steroid hormone responses, which can impact the tissue’s functional and physiologic roles of pregnancy establishment and maintenance, tissue homeostasis after menses and regeneration, and endometrial disorders. With regard to endometriosis, aberrant promoter methylation of several genes whose products are critical for the normal endometrial P4 response (e.g., PGRB, HOXA10, ESR2, SF1) has been reported.
in eutopic endometrium and ectopic disease in the pelvis of women with endometriosis and in animal models of the disease, with resulting resistance to P4 action [15–17].

Herein, we investigated the global DNA methylation of proliferative, early secretory, and midsecretory endometrium of women with severe endometriosis. We also examined the association of DNA methylation changes with gene expression assessed by microarray analysis of the same samples. The results show that the DNA methylome across the cycle differs from that in controls, with most abnormalities observed in the early and midsecretory (progesterone-dominant) phases and involving mostly CGIs. Elucidating hormone-dependent epigenetic states required for normal nuclear receptor function and discovering epigenetic reprogramming events contributing to hormone resistance could have broader implications for other hormone-dependent disorders.

**MATERIALS AND METHODS**

**Sample Collection and Processing**

This study was approved by the Committee on Human Research of the University of California, San Francisco (UCSF). Only eutopic endometrial samples were studied. All patients undergoing surgery for endometriosis-related pain and/or infertility gave written informed consent through the UCSF NIH Human Endometrial Tissue and DNA Bank [18] (Supplemental Table S1; Supplemental Data are available online at www.biolreprod.org). Seventeen eutopic endometrial tissue samples were from patients with severe endometriosis (stage IV; n = 4 in proliferative [PE], n = 7 in early secretory [ESE], and n = 6 in midsecretory [MSE] phases). Disease severity was determined according to the revised American Society for Reproductive Medicine (ASRM) staging [19]. Control samples (no endometriosis; n = 6 PE, n = 5 ESE, and n = 5 MSE) were from a previous study [14]. Participants were nonsmokers (two exceptions), were not pregnant, and had had no hormonal treatments within 3 months prior to sample acquisition. Tissue procurement, processing, and storage were conducted as described previously [14]. Menstrual cycle phase was determined by histology, serum levels of E2 and P4, and/or unsupervised principal component analysis and independent hierarchical clustering analysis of microarray gene expression data [10, 14]. Genomic DNA was extracted and stored as described previously [14].

**DNA Methylation Analysis**

Genomic DNA bisulfite conversion and quality controls (QCs) were conducted as described previously [14, 20]. Briefly, bisulfite conversion was done using the Zymo EZ-96 DNA methylation Kit (Zymo Research, Irvine, CA) based on the manufacturer’s protocol. The QCs included a panel of MethyLight reactions [20] to assess the completeness of the bisulfite conversion and the amount of bisulfite-converted DNA. The quantity and integrity of the sample DNA after bisulfite conversion were determined by a bisulfite-dependent, methylation-independent MethyLight reaction of a multiplex sequence, which is well distributed across the genome [20]. The completeness of bisulfite treatment was assessed using a panel of bisulfite-independent primers with variable bisulfite-dependent probes reflecting various degrees of bisulfite conversion in the sample, including full conversion (100% conversion), no conversion (0% conversion), and partial conversion (50% conversion) [20]. All samples, except two (one ESE and one MSE), passed all QCs and were further assayed by the quantitative Illumina HumanMethylation27K platform (Illumina, San Diego, CA) based on the manufacturer’s specifications, as described previously [14]. The Illumina 27K platform interrogates DNA methylation levels at 27578 CpG sites corresponding to 14475 protein-coding and 110 micro-RNA coding genes.

DNA methylation values were scored as β values (ratio of methylated signal over total fluorescent signal), ranging from 0 to 1 (from no methylation to complete methylation, respectively). DNA methylation measurement quality for each probe in each sample was assessed by the detection P value, calculated based on the difference in its signal intensity compared with a set of 16 negative control probes; only probes with detection P values <0.05 (i.e., statistically significant differences from background) were retained for further analysis, and those with P > 0.05 were marked as “missing” and were excluded from further analyses. For each sample, the probe dropout rate was calculated as the percent of total number of probes with “missing” values from the total number of platform probes (27578). Using stringent criteria, another ESE sample with highest “missing” values (3.5% dropout rate) was excluded from subsequent analyses. The dropout rates for all other samples were <1% of total platform probes (0%-0.8%). Thus, the final endometriosis samples included in the analyses were as follows: n = 4 PE, n = 5 ESE, and n = 5 MSE.

**Global DNA Methylation Profile in Endometriosis**

Probes with a “missing” value in more than one sample (n = 116) were removed. The remaining 27462 probes were assessed for global DNA methylation profiles and patterns across the menstrual cycle in endometriosis. Hypermethylated CpG sites are defined as having β value >0.8; hypomethylated CpG sites, β value <0.2; and intermediate methylated CpG sites, 0.2 ≤ β ≤ 0.8. The state of methylation for each probe was determined for each cycle phase and then across other phases to determine if and which probes were hypermethylated, hypomethylated, or intermediately methylated in either all or some of the phases. The CpG sites hypomethylated or hypermethylated in all phases, intermediately methylated, or hypomethylated in some but not all phases were next compared to the corresponding groups in control endometrium to determine the extent of similarities and differences in profiles, patterns, and loci in disease versus control.

**Differentially Methylated CpG Sites in Cycle Phases**

Median β values for each probe in each phase were calculated, and median β value differences were used to identify differentially methylated CpG sites between cycle phases. Probes with more than one missing value in each group were excluded from analysis. Cycle phase-specific median β value differences between endometriosis and controls were derived by subtracting the median β value of each probe in a specific phase from the corresponding value in endometriosis, resulting in the following comparisons: PEendo versus PEcontrol, ESEendo versus ESEcontrol, and MSEendo versus MSEcontrol. Also, comparisons across phases of the cycle were conducted for endometriosis endometrial samples (ESEendo vs. PEendo, MSEendo vs. ESEendo, and MSEendo vs. PEendo). Differentially methylated CpG sites in the different phase comparisons in disease were also compared to previously reported changes across the cycle in control endometrium. In addition, all endometriosis samples were compared to all control samples to investigate phase-independent differentially methylated loci between disease and controls. For all comparisons, probes were considered differentially methylated only with Δβ >0.136 (detectable differences of 95% confidence) [21]. As previously determined [21], smaller differences in β values may be unreliable because of background noise and platform variability.

**RNA Isolation and Gene Expression Microarray**

Portions of the same tissue samples were processed for total RNA isolation with DNase treatment using an RNase Bu Plus Kit (Valencia, CA), according to the manufacturer’s specification. RNA quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA sample preparation, quality assessment, and hybridization to Affymetrix HU133 Plus 2.0 gene expression arrays (>38500 genes; Affymetrix, Santa Clara, CA) was conducted as previously described [14].

**Comparison of DNA Methylation and Gene Expression Data**

The raw .cel expression data were GCRMA normalized using GeneSpring GX 12.0 software (Agilent Technologies). The transcription unit identifier was used to match corresponding probes from the two platforms using R (http://www.r-project.org). Every DNA methylation probe for a given locus was compared to all transcripts of the corresponding locus. DNA methylation association with gene expression was investigated using the Spearman correlation.

**Gene Functional and Biological Classification Analyses**

The DAVID database [22] was used to investigate functional classification of genes that were differentially methylated in the same phases of disease versus control and across the cycle in disease. Furthermore, gene lists were cross-referenced to published data of transcriptome differences in endometriosis in women [8, 23–26] and a baboon model of the disease [27]. In addition, an extensive literature review was conducted of the genes’ biological characteristics and functions to evaluate genes and/or gene groups potentially important in the disease’s pathophysiology.
RESULTS

DNA Methylation Profiles of Eutopic Endometrium in Endometriosis During the Menstrual Cycle

DNA methylation profiles of endometrium of patients with severe endometriosis were investigated for: 1) global profiles and patterns across the cycle; 2) phase-specific differences versus controls; and 3) differences across the cycle in disease.

Global DNA Methylation Profiles and Patterns Across the Menstrual Cycle

To better understand the global changes in profiles and patterns of DNA methylation across the phases of the cycle in endometriosis, we assessed the frequency of differentially methylated loci, their state of loss or gain of methylation, and their genomic and chromosomal distribution, as well as their association within or outside CGIs. In endometrium from women with endometriosis, most CpG sites (61%; 16,765 CpG sites) were hypomethylated (β-value < 0.2) in all cycle phases, whereas a small percentage (10%; 2,803 CpG sites) were hypermethylated (β-value > 0.8) in all phases or were hypomethylated (β < 0.2: 1.8%; 489 CpG sites) or hypermethylated (β > 0.8: 2.2%; 610 CpG sites) in some but not all phases (Fig. 1a). This general pattern is very similar to that of control endometrium (61%; white) (Fig. 1b) and similar to somatic tissue global DNA methylation patterns (i.e., hypermethylation involving non-CGI CpG sites and a small fraction of CGIs, with most CGIs remaining hypomethylated) [13, 28].

Most of the loci that are hypomethylated (β value < 0.2) or hypermethylated (β value >0.8) in all cycle phases (i.e., phase independent) in disease were also hypomethylated or hypermethylated in all cycle phases in controls (disease independent). The hypomethylated group had 97% (16,538 CpG sites) and the hypermethylated group had 95% (2,668 CpG sites) in endometriosis (Fig. 1a). Global DNA methylation pattern of eutopic endometrium in endometriosis and its comparison with normal eutopic endometrium based on both methylation levels and CpG site location within or outside CpG islands.

FIG. 1. Global DNA methylation pattern of eutopic endometrium in endometriosis and its comparison with normal eutopic endometrium based on both methylation levels and CpG site location within or outside CpG islands. 

Legend: 
- **a.** Endometriosis 
- **b.** Control 
- **c.** In non-CGI [968 (6%)] 
  - In CGI [15,797 (94%)] 
  - Total [16,765 (61%)] 
- **d.** In non-CGI [1,006 (6%)] 
  - In CGI [15,724 (94%)] 
  - Total [16,730 (61%)] 
- **e.** In non-CGI [1,863 (66.5%)] 
  - In CGI [940 (33.5%)] 
  - Total [2,803 (10%)] 
- **f.** In non-CGI [1,885 (65%)] 
  - In CGI [998 (35%)] 
  - Total [2,881 (10%)]
common with control endometrium (Supplemental Table S2). Therefore, these loci could be regarded as both cycle and disease independent endometrial hypomethylated/hypermethylated loci. However, their endometrial uniqueness remains to be determined upon extensive comparisons to other tissues. The phase- and disease-independent hypomethylated group comprised 60% of the platform loci and therefore includes genes across the genome and with various functions. The hypermethylated group, with a smaller number of loci (9.7% of the platform), included genes in signaling and signal peptides, glycoproteins, defense response, plasma membrane, spermogenesis, sexual reproduction, and gamete generation, among others (Supplemental Table S2).

The state of methylation differences between disease and control for the remaining loci, and with variable or intermediate methylation levels were further investigated in the context of phase-specific differences in disease versus controls as well as DNA methylation changes across the cycle in disease (see below).

**Phase-Specific Differentially Methylated Loci in Endometriosis Versus Controls**

Phase-specific DNA methylation profiles were compared (PEEndo vs. PEControl, ESEEndo vs. ESEControl, MSEEndo vs. MSEC ontrol, Table 2) and revealed phase-specific differences in disease versus controls in the: 1) DNA methylation profiles, 2) frequencies of gain or loss of methylation, and 3) involvement of CGIs—suggestive of the interaction of the methylome with varying hormonal milieu. These are specifically addressed below.

**DNA methylation profiles in endometriosis.** The largest differences in the number of differentially methylated CpG sites were in the Pδ-dominant midsecretory phase—that is, in MSE versus MSEC ontrol (137 CpG sites, corresponding to 125 loci; Fig. 2, right), followed by PEEndo versus PEControl (58 CpG sites, corresponding to 58 loci; Fig. 2, left) and ESEEndo versus ESEC ontrol (39 CpG sites, corresponding to 36 loci; Fig. 2, middle). A representative selection of phase-specific differentially methylated CpG sites in disease versus control with biological relevance in endometrium and endometriosis is shown in Table 1 (see Supplemental Table S3 for full gene list). Differentially methylated CpG sites in disease versus control were mostly unique in each phase comparison (i.e., phase specific), with only a small number of differentially methylated loci in common between phases (Table 2).

**Gain and loss of methylation.** Phase-specific differential methylation in disease versus control included both gain and loss of methylation (Fig. 3), but patterns of methylation gain versus loss were different in the proliferative and secretory phases (Fig. 3). In PEEndo versus PEControl most (64%) of the differentially methylated CpG sites gained methylation in disease (Fig. 3a), whereas in ESEEndo versus ESEC ontrol, most (79%) lost methylation in disease (Fig. 3b). In MSEEndo versus MSEC ontrol, the frequencies of gain and loss of methylation in disease were similar (Fig. 3c). Relevant to endometrium and endometriosis, some of the genes more methylated in PEEndo versus PEControl included GSTM1, GSTM5, HOX5, and FAIM2; those less methylated in PEEndo versus PEControl included TAFID, IL17B, and TRPM1. Gain of methylation in MSEEndo versus MSEC ontrol included genes such as TOB1, VNNT1, BDH2, and NPSR1, and loss of methylation in MSEEndo versus ESEControl included FZD2, HOXA9, HOXD12, ALG13, and CLEC7A. In MSEEndo versus MSEC ontrol, MPP7, PLEK, and DG5 were located outside CGIs, and COL12A1, KCNE4, EDNRB, HOXD12, LAMA3, and HOXA9 were located inside CGIs.

**Phase-independent differentially methylated loci in disease versus control.** The DNA methylene of all disease samples was compared to that of all controls across all phases, and only three CpG sites were differentially methylated, corresponding to three genes: RPF2, more methylated in controls than disease and within a CGI; and PER1 and FAM181A, both more methylated in disease than controls and within and outside a CGI, respectively (Table 2). Because of the changes of DNA methylome throughout the phases of the cycle in both disease and control, this small number was not unanticipated. Several other loci, including HOXA9, HOXD12, IRX2, NKKX6-2, CYP7B1, ALG13, MYO3A, and FZD2, all within CGIs, were less methylated in disease versus controls in both ESE and MSE phases (Table 2).

**Association of changes in DNA methylation with gene expression.** To investigate if phase-specific differentially methylated CpG sites in disease versus control endometrium were associated with changes in gene expression, we evaluated the gene expression profile of a portion of the same tissue samples by whole-genome microarray analysis (see Supplemental Fig. S1 for patterns of gene expression changes independent of DNA methylation). We also assessed the relationship of DNA methylation levels of each locus to the gene expression of its corresponding transcripts from the microarrays (see Materials and Methods). Because of the nonlinear nature of the relationship of DNA methylation with gene expression [29], the Spearman correlation was used.

Phase-specific differentially methylated CpG sites between disease and control showed both positive and negative associations with changes in gene expression (Table 3). This is expected because DNA methylation at gene promoters, either at CGIs or non-CGIs, is usually associated with decreased gene expression of its corresponding transcripts from the microarrays (see Materials and Methods). Because of the nonlinear nature of the relationship of DNA methylation with gene expression [29], the Spearman correlation was used.
expression, whereas gene body methylation is usually associated with increased expression [11, 30, 31]. Also, DNA methylation at CGIs usually shows both negative and positive association with gene expression, although non-CGI methylation is usually negatively associated with gene expression [13]. CpG sites differentially methylated in disease versus controls in PE and MSE showed similar strengths of positive and negative association with gene expression, with stronger negative association in \( PE_{	ext{Endo}} \) versus \( PE_{	ext{Control}} \) (\( +\rho = 0.24, -\rho = -0.30 \)) and equal positive and negative association in \( MSE_{	ext{Endo}} \) versus \( MSE_{	ext{Control}} \) (\( +\rho = 0.28, -\rho = -0.28 \)). Further analysis of CpG context showed that for CpG sites located within CGIs, almost equal numbers had positive or negative association with gene expression, whereas most CpG sites located outside CGIs showed negative association with gene expression. However, in \( ESE_{	ext{Endo}} \) versus \( ESE_{	ext{Control}} \) most
| Illumina probe ID  | Gene symbol     | Product   | CGI  | Median $\beta$ value difference |
|-------------------|-----------------|-----------|------|-------------------------------|
| cg01980637        | PER1            | Period 1  | Yes  | 0.14                          |
| cg00929855        | HSPA1A          | Heat shock 70-kDa protein 1A | Yes  | 0.17                          |
| cg17901463        | GSTM1           | Glutathione S-transferase M1 isoform 1 | Yes  | 0.25                          |
| cg2248486         | HOXA5           | Homeobox A5 | Yes  | 0.17                          |
| cg24464875        | TCEAL3          | Transcription elongation factor A (SII)-like 3 | Yes  | 0.14                          |
| cg04987894        | GSTM5           | Glutathione S-transferase M5 | Yes  | 0.17                          |
| cg08717396        | HIST1H2AG       | H2A histone family; member P | Yes  | 0.15                          |
| cg10143146        | COL11A2         | Collagen; type XI; alpha 2 isoform 3 preproprotein | Yes  | 0.16                          |
| cg27188703        | FAIM2           | Fas apoptotic inhibitory molecule 2 | Yes  | 0.17                          |
| cg18241160        | CDK11A          | Cell division cycle 2-like 2 isoform 2 | Yes  | 0.16                          |
| cg13915726        | DUSP9           | Dual-specific phosphatase 9 | Yes  | 0.25                          |
| cg07404485        | PON1            | Paraoxonase 1 | No   | 0.16                          |
| cg15518950        | TMEM171         | Proline-rich protein PRP2 | Yes  | 0.20                          |
| cg18085517        | TRPM1           | Transient receptor potential cation channel; subfamily M; member 1 | No   | -0.17                         |
| cg13603171        | MOXD1           | Monoxygenase; DBH-like 1 isoform 2 | Yes  | -0.14                         |
| cg01963696        | ELANE           | Elastase 2; neutrophil preproprotein | Yes  | -0.15                         |
| cg21296602        | TAF1D           | TATA-box-binding protein associated factor, RNA polymerase I subunit D | No   | -0.15                         |
| cg25141490        | IL17B           | Interleukin 17B precursor | No   | -0.14                         |
| cg24315815        | PLSCR4          | Phospholipid scramblase 4 | Yes  | -0.15                         |
| cg15928132        | CCKAR           | Cholecystokinin A receptor | No   | -0.14                         |
| cg06506864        | NPSR1           | G protein-coupled receptor 154 isoform B | No   | -0.14                         |
| cg1531099         | LCE1D           | Late cornified envelope 1D | No   | 0.20                          |
| cg06506864        | NPSR1           | G protein-coupled receptor 154 isoform B | No   | 0.17                          |
| cg02214188        | BDH2            | 3-Hydroxybutyrate dehydrogenase; type 2 | No   | 0.16                          |
| cg00852964        | VNN1            | Vanin 1 precursor | No   | 0.15                          |
| cg14494812        | TOB1            | Transducer of ERBB2; 1 | No   | 0.15                          |
| cg22373097        | KRTAP21-1       | Keratin-associated protein 21-1 | No   | -0.17                         |
| cg22825487        | VNN5            | Vanin 3 isoform 2 precursor | No   | -0.17                         |
| cg05213296        | RPF2            | Brix domain containing 1 | Yes  | -0.14                         |
| cg03874199        | HOXD12          | Homeobox D12 | Yes  | -0.14                         |
| cg25228126        | FZD2            | Frizzled 2 | Yes  | -0.15                         |
| cg02757432        | GPR26           | G protein-coupled receptor 26 | Yes  | -0.15                         |
| cg19963797        | ALG13           | Glycosyltransferase 28 domain containing 1 | Yes  | -0.15                         |
| cg13152535        | CLEC11A         | Stem cell growth factor precursor | Yes  | -0.16                         |
| cg26521404        | HOXA9           | Homeobox protein A9 isoform a | Yes  | -0.16                         |
| cg10146929        | HIST1H1A        | H1 histone family; member 1 | Yes  | -0.16                         |
| cg15433631        | IRX2            | Iroquois homeobox protein 2 | Yes  | -0.16                         |
| cg08441806        | NKX6-2          | NK6 transcription factor related; locus 2 | No   | -0.17                         |
| cg04418492        | CYPTB1          | Cytochrome P450; family 7; subfamily B; polypeptide | No   | -0.18                         |
| cg20588069        | MSX1            | Msx homeobox homolog 1 | Yes  | -0.19                         |
| cg23873703        | KCNAB1          | Potassium voltage-gated channel; shaker-related subfamily; beta member 1 isoform 2 | No   | -0.19                         |
| cg23771603        | MYO3A           | Myosin IIIA | Yes  | -0.20                         |
| cg16192575        | PTPN20B         | Protein tyrosine phosphatase; nonreceptor type 20 | Yes  | -0.23                         |
| cg14528319        | GIPC1           | Regulator of G-protein signaling 19 interacting protein 1 isoform 1 | No   | 0.17                          |
| cg12743398        | SULT1A2         | Sulforotransferase family; cytosolic; 1A; phenol-prefering; member 2 | No   | 0.16                          |
| cg25141490        | IL17B           | Interleukin 17B precursor | No   | 0.16                          |
| cg12069042        | PLXNB1          | Plexin B1 | No   | 0.15                          |
| cg20530056        | IKBKE           | IKB-related kinase epsilon | No   | 0.15                          |
differentially methylated CpGs were positively associated with gene expression and the positive association was stronger \((+p = 0.40, -p = -0.32; \text{Table } 3\)) \(\text{Table 3}\). CpG island analysis showed that most differentially methylated CpG sites within CGIs were positively associated with gene expression, whereas CpG sites outside CGIs had equal numbers positively or negatively associated with gene expression. These differences between cycle phases suggest complex recruitment of sequence-specific epigenetic machinery in response to different hormonal milieu in the endometrium of endometriosis patients.

Loci with moderate/high correlation between DNA methylation and gene expression. Differentially methylated

| Illumina probe ID | Gene symbol | Product | CGI | Median \(\beta\) value difference |
|-------------------|-------------|---------|-----|---------------------------------|
| cg23349242        | SUSD2       | Sushi domain containing 2 | No  | 0.14                            |
| cg04052038        | CLDN8       | Claudin 8 | No  | 0.14                            |
| cg15784615        | LTBR        | Lyphotoxin beta receptor | No  | 0.14                            |
| cg06339706        | PLEKH4A     | Pleckstrin homology domain containing; family A | No  | 0.14                            |
| cg09837648        | PLXNB1      | Plexin B | No  | 0.14                            |
| cg15720535        | AGPAT2      | 1-Acylglycerol-3-phosphate O-acyltransferase 2 | Yes | 0.26                            |
| cg12120741        | EDNRB       | Endothelin receptor type B isoform 2 | Yes | 0.21                            |
| cg13439730        | PRSS8       | Prostasin preproprotein | Yes | 0.21                            |
| cg20837775        | SERPINB5    | Serine (or cysteine) proteinase inhibitor; clade B (ovalbumin); member 5 | Yes | 0.20                            |
| cg04106785        | CDK5R1      | cyclin-dependent kinase 5; regulatory subunit 1 | Yes | 0.18                            |
| cg09835085        | KCNE4       | Potassium voltage-gated channel; Isk-related family; member 4 | Yes | 0.18                            |
| cg10313633        | TP53I11     | p53-induced protein | Yes | 0.17                            |
| cg12707353        | C4orf23     | Hypothetical protein LOC152992 isoform 2 | Yes | 0.16                            |
| cg24183173        | BCROR       | BCL-6 interacting corepressor 2 | Yes | 0.15                            |
| cg19264571        | APCDD1      | Adenomatosis polyposis coli down-regulated 1 | Yes | 0.14                            |
| cg23165341        | DAPK2       | Death-associated protein kinase 2 | Yes | 0.14                            |
| cg17864555        | SCAND2      | SCAN domain-containing protein 2 isoform 1 | Yes | 0.14                            |
| cg02794695        | SLA         | Src-like-adaptor | No  | -0.14                           |
| cg01361777        | DLG5        | Discs large homolog 5 | No  | -0.14                           |
| cg04872689        | PLEK        | Pleckstrin | No  | -0.21                           |
| cg03874199        | HOXD12      | Homeobox D12 | No  | -0.14                           |
| cg24199834        | POU4F2      | POU domain; class 4; transcription factor 2 | Yes | 0.14                            |
| cg00662775        | TCEAL4      | Transcription elongation factor A (SII)-like 4 | Yes | 0.14                            |
| cg24101578        | CDH22       | Cadherin 22 precursor | Yes | 0.14                            |
| cg19352038        | PAX3        | Paired box gene 3 isoform PAX3 | Yes | 0.14                            |
| cg17347634        | CYP7B1      | Cytochrome P450; family 7; subfamily B; polypeptide 1 | Yes | 0.14                            |
| cg22815110        | FOXD3       | Forkhead box D3 | Yes | 0.14                            |
| cg01009664        | THF        | Thyrotropin-releasing hormone | Yes | 0.14                            |
| cg19205533        | RERG        | RAS-like; estrogen-regulated; growth inhibitor | Yes | 0.15                            |
| cg18905252        | CFC1        | Cryptic. | Yes | 0.15                            |
| cg25228126        | FZD2        | Frizzled 2 | Yes | 0.15                            |
| cg14625095        | HIST1H1A    | H1 histone family; member 1 | Yes | 0.15                            |
| cg14894144        | LAMA3       | Laminin alpha 3 subunit isoform 2 | Yes | 0.18                            |
| cg15992730        | GDF3        | Growth differentiation factor 3 precursor | Yes | 0.15                            |
| cg16636586        | BST2        | Bone marrow stromal cell antigen 2 | Yes | 0.15                            |
| cg08441806        | NFKX6-2     | NFK6 transcription factor related; locus 2 | Yes | 0.15                            |
| cg08047457        | RASSF1      | Ras association domain family 1 isoform A | Yes | 0.16                            |
| cg25094569        | WT1         | Wilms tumor 1 isoform C | Yes | 0.16                            |
| cg17965019        | HIST1H3J    | H3 histone family; member J | Yes | 0.16                            |
| cg24628744        | H2AFY       | H2A histone family; member Y isoform 2 | Yes | 0.16                            |
| cg22709192        | HOXC11      | Homeobox C11 | Yes | 0.17                            |
| cg1343482         | GAT4        | GATA binding protein 4 | Yes | 0.17                            |
| cg14951292        | HMOX2       | Heme oxygenase (decyclizing) 2 | Yes | 0.17                            |
| cg00842351        | TIP2        | Tjight junction protein 2 (zona occludens 2) isoform 2 | Yes | 0.18                            |
| cg10210238        | CDKN2B      | Cyclin-dependent kinase inhibitor 2B isoform 2 | Yes | 0.18                            |
| cg01683883        | CMTM2       | Chemokine-like factor superfamily 2 | Yes | 0.18                            |
| cg11233198        | CDH8        | Cadherin 8; type 2 preproprotein | Yes | 0.18                            |
| cg23771603        | MYO3A       | Myosin IIIA | Yes | 0.18                            |
| cg23290344        | NEFM        | Neurofilament 3 (150-kDa medium) | Yes | 0.20                            |
| cg11381846        | HOX9A       | Hox homolog protein A9 isoform a | Yes | 0.21                            |
| cg27652350        | ALDH1A3     | Aldehyde dehydrogenase 1A3 | Yes | 0.24                            |
| cg07533148        | TRIM58      | Tripartite motif-containing 58 | Yes | 0.25                            |
| cg05213296        | RFP2        | Brix domain-containing 1 | Yes | 0.31                            |
TABLE 2. Common loci differentially methylated between phases of disease versus control.a

| Gene symbol | CGI | Gene | Comparison | Median $\beta$ value difference | Median $\beta$ value difference |
|-------------|-----|-------|------------|-------------------------------|-------------------------------|
| cg08441170  | Yes | RPF2  | Endo > Control | 0.15                          | 0.16                          |
| cg01980637  | Yes | PER1  | Control > Endo | 0.16                          | 0.17                          |
| cg20022541  | No  | FAM181A | Endo > Control | -0.17                         | -0.17                         |
| cg06506864  | No  | NPSR1 | ESE Endo > MSE Control | 0.20                          | 0.16                          |
| cg03874199  | Yes | HOXD12 | ESE Endo > ESE Control | 0.14                          | 0.17                          |
| cg19963797  | Yes | ALG13 | ESE Control > MSE Control | 0.15                          | 0.16                          |
| cg25228126  | Yes | FZD2  | ESE Control > MSE Control | 0.15                          | 0.15                          |
| cg08441806  | Yes | NKX6-2 | ESE Endo > MSE Endo | 0.14                          | 0.15                          |
| cg15433631  | Yes | IRX2  | MSE Endo > MSE Control | 0.14                          | 0.15                          |
| cg15431806  | Yes | CYP7B1 | ESE Endo > MSE Endo | 0.14                          | 0.18                          |
| cg23771603  | Yes | MYO3A | ESE Endo > MSE Endo | 0.16                          | 0.18                          |

a Cont, control.

**Differentially Methylated Loci Across the Cycle in Endometriosis**

Phase comparison analyses (i.e., different hormonal milieu, ESEEndo vs. PEEndo, MSEEndo vs. PEEndo, and MSEEndo vs. ESEControl, Fig. 4) were conducted to investigate hormone dependence/phase differences and potential abnormalities in the setting of endometriosis.

**Cycle-phase methylation differences.** The largest differences in DNA methylation were between MSEEndo versus ESEEndo (100 CpG sites, corresponding to 96 loci; Fig. 4, right), representing maximum P4 action (MSE) and rising P4 levels (ESE), followed by differences between MSEEndo (maximum P4) versus PEEndo (maximum E2; 81 CpG sites, corresponding to 79 loci; Fig. 4, middle). The fewest differences were observed in ESEEndo versus PEEndo (34 CpG sites, corresponding to 32 loci; Fig. 4, left; Supplemental Table S4, full gene list). For example, KRT19, CDK11A, and KCNC3 were more methylated in PEEndo versus ESEEndo; RUNX3, CKDN2B, and PLEK were more methylated in ESEEndo versus MSEEndo; and CASP8, ALDH1A3, and PTPRC were more methylated in PEEndo versus MSEEndo.

We previously found that normal endometrium displays changes between cycle phases [14]. Herein, the (above) cycle-specific differentially methylated loci were compared with those of normal endometrium. In ESE versus PE, there were no common differentially methylated loci between disease (34 CpG sites) and normal (27 CpG sites) endometrium (Supplemental Table S5). In MSE versus ESE, there was only one locus, FAM181A, in common between differentially methylated loci in disease (100 CpG sites) and normal (22 CpG sites; Supplemental Table S5) endometrium. FAM181A was more methylated in MSE than ESE in both disease and normal endometrium. In MSE versus PE, five loci are in common, but only one locus, FAM181A, showed the same direction of methylation change in both disease and normal, and the other four loci showed the opposite direction of differential methylation in MSE versus PE in disease and normal endometrium (Supplemental Table S5): FAM181A was more methylated in MSE than PE in both disease and normal endometrium; TAF1D and C21orf128 were less methylated in MSE than PE in normal, but more methylated in MSE than PE in disease; and PM20DA and ATP8A2 were more methylated in MSE than PE in normal, with the opposite pattern in disease. Aside from little overlap in cycle-phase changes with normal
endometrium, the number of differently methylated loci in cycle-phase comparisons was larger in disease than in normal. Furthermore, in normal endometrium, the least number of differences were between MSE versus ESE (22 CpG sites), and the most differences were seen in MSE versus PE (peak 2P vs. peak E2; 66 CpG sites) [14], which is different from disease, with most differences in MSE versus ESE (100 CpG sites) and least between PE and ESE (34 CpG sites). These differences, as well as methylation gain/loss pattern and CGI involvements (see below), suggest broad abnormalities in endometrial DNA methylation across the cycle in endometriosis.

**Gain and loss of methylation.** When comparing methylation changes between the proliferative with either of the secretory subphases (Fig. 5, a and c), most CpG sites showed loss of methylation (85% PE > ESE and 70% PE > MSE). However, within the secretory phase, with the highest differences in DNA methylation between ESE and MSE, loss and gain were observed, with a higher percentage (62%) more methylated in MSE versus ESE (Fig. 5b). In control endometrium, the frequency of gain and loss of methylation is similar in all three phase comparisons [14]. This differs from disease, with more loss than gain of methylation in ESE versus PE (85% loss vs. 15% gain) and MSE versus PE (70% loss, 30% gain), and more gain than loss of methylation in MSE versus ESE (62% gain, 38% loss).

**Association with CGIs.** In general, most differentially methylated CpG sites across the cycle in endometriosis were within CGIs (Fig. 4). Differentially methylated CpG sites in ESEEndo versus PEEndo were mostly located in CGIs for loci, with either gain or loss of methylation in ESEEndo compared with PEEndo (Fig. 5a). In MSEEndo versus PEEndo and in MSEEndo versus ESEEndo, most CpG sites with loss of methylation in MSE were in CGIs (Fig. 5, b and c), whereas gain of methylation in MSE compared with ESEEndo or PEEndo occurred with similar frequency within or outside CGIs (Fig. 5, b and c). These observations differ from control endometrium, wherein methylation changes involved both CGIs and non-CGIs with similar frequencies in all comparisons. Also, in controls, CpG sites that had gain of methylation in MSE compared with PE were mostly located within a CGI, and CpG sites that showed loss of methylation in MSE compared with PE were mostly located outside CGIs [14]—opposite the pattern observed in disease herein.

**Association of changes in DNA methylation with gene expression.** Differentially methylated CpG sites in ESEEndo versus PEEndo, MSEEndo versus ESEEndo, and MSEEndo versus PEEndo showed positive and negative associations with gene expression changes, with positive association frequency slightly higher in all three comparisons (Table 3). Also, CpG sites within CGIs showed both positive and negative associations with gene expression, with a higher frequency of positive association. For CpG sites located outside CGIs, higher frequency of negative association was observed in MSEEndo versus PEEndo, whereas equal frequency of positive and negative association was observed in MSEEndo versus ESEEndo. There were no CpG sites located outside CGIs in MSEEndo versus PEEndo associated with gene expression (Table 3).

**TABLE 3. Relationship between location within or outside CGI with positive or negative association between DNA methylation and gene expression as measured by Spearman rho.**

| Comparison            | Total rho | Mean positive rho | Mean negative rho | % In CGI | % Out CGI | % In CGI positive | % In CGI negative | % Out CGI positive | % Out CGI negative |
|-----------------------|-----------|-------------------|-------------------|----------|----------|-------------------|-------------------|--------------------|--------------------|
| PE endo versus PE control | −0.06     | 0.24              | −0.30             | 73.9     | 26.1     | 51.2              | 48.8              | 24.1               | 75.9               |
| ESE endo versus ESE control | 0.13     | 0.40              | −0.32             | 80.6     | 19.4     | 66.7              | 33.3              | 46.2               | 53.8               |
| MSE endo versus MSE control | 0.01     | 0.28              | −0.28             | 79.3     | 20.7     | 54.0              | 46.0              | 41.1               | 58.9               |
| ESE endo versus PE endo | 0.12     | 0.44              | −0.30             | 96.8     | 3.2      | 59.0              | 41.0              |                    |                    |
| MSE endo versus ESE endo | 0.04     | 0.38              | −0.38             | 63.2     | 36.8     | 58.3              | 41.7              | 52.2               | 47.8               |
| MSE endo versus PE endo | 0.06     | 0.36              | −0.32             | 66.5     | 33.5     | 67.3              | 32.7              | 33.3               | 66.7               |
**TABLE 4.** Select differentially methylated loci associated with gene expression in phases of disease versus control.

| Illumina probe ID | Gene symbol | Product | CGI | Median β value difference | Affymetrix probe set ID | Fold change | Spearman rho |
|-------------------|-------------|---------|-----|----------------------------|-------------------------|-------------|--------------|
| Endo PE versus control PE |
| cg21296602 | TAF1D | TATA-box-binding protein-associated factor, RNA polymerase I, subunit D | No | —0.15 | 221580_s_at | 1.8 | —0.4 |
| cg04987894 | GSTM5 | Glutathione S-transferase M5 | Yes | 0.17 | 205752_s_at | —1.2 | —0.5 |
| cg04995717 | TEK | TEK tyrosine kinase; endothelial | No | 0.16 | 206702_s_at | —1.4 | —0.1 |
| cg17901463 | GSTM1 | Glutathione S-transferase M1 | Yes | 0.25 | 215333_s_at | —1.5 | —0.6 |
| cg11965370 | NTM | Neurotrin | No | —0.16 | 227566_at | 2.0 | 0.6 |
| cg27188703 | FAIM2 | Fas apoptotic inhibitory molecule 2 | Yes | 0.17 | 203619_s_at | —2.2 | —0.6 |
| cg02248486 | HOXA5 | Homeobox A5 | Yes | 0.17 | 213844_at | —3.3 | —0.8 |
| cg12061236 | AKAP12 | A-kinase anchor protein 12 | Yes | 0.14 | 227530_at | —1.9 | —0.7 |
| Endo ESE versus control ESE |
| cg23873703 | KCNA1 | Potassium voltage-gated channel; shaker-related subfamily; beta member 1 isoform 2 | Yes | —0.19 | 210471_s_at | 1.9 | —0.5 |
| cg25228126 | FZD2 | Frizzled 2 | Yes | —0.15 | 210220_at | —4.0 | 0.3 |
| cg19963797 | ALG1 | Glycosyltransferase 28 domain-containing 1 | Yes | —0.15 | 219013_s_at | 4.3 | —0.8 |
| cg20616414 | WNK2 | WNK lysine-deficient protein kinase 2 | Yes | —0.15 | 228208_at | 1.8 | —0.1 |
| cg14494812 | TOB1 | Transducer of ERBB2; 1 | No | 0.15 | 228834_at | 6.7 | 0.9 |
| cg21561173 | C2orf81 | Hypothetical protein LOC114035 | No | 0.15 | 1569067_s_at | 7.0 | 0.7 |
| cg21561173 | C2orf81 | Hypothetical protein LOC114035 | No | 0.15 | 241233_s_at | 1.8 | 0.8 |
| cg02214188 | BDH2 | 3-Hydroxybutyrate dehydrogenase; type 2 | No | 0.16 | 235155_at | 1.6 | 0.4 |
| Endo MSE versus control MSE |
| cg04872689 | PLEK | Pleckstrin | No | —0.21 | 203471_s_at | 1.6 | —0.8 |
| cg01354473 | HOXA5 | Homeobox protein A9 isoform a | Yes | —0.21 | 214651_s_at | 1.5 | —0.3 |
| cg14904144 | LAMC2 | Laminin alpha 3 subunit isoform 2 | No | —0.18 | 203726_s_at | 3.5 | —0.8 |
| cg16363586 | BST2 | Bone marrow stromal cell antigen 2 | Yes | —0.15 | 201641_at | 1.3 | —0.7 |
| cg01361777 | DLG5 | Discs large homolog 5 | No | —0.14 | 201681_s_at | —1.6 | 0.9 |
| cg01361777 | DLG5 | Discs large homolog 5 | No | —0.14 | 214649_at | —2.7 | 0.8 |
| cg02794695 | SLA | Src-like adaptor | No | —0.14 | 203761_at | 2.1 | —0.8 |
| cg26780404 | COL12A1 | Alpha 1 type XII collagen short isoform precursor | Yes | 0.14 | 225664_at | 2.6 | 0.5 |
| cg01980637 | PER1 | Period 1 | Yes | 0.15 | 36829_at | —2.2 | —0.6 |
| cg21663431 | SLC44A2 | CTL2 protein | Yes | 0.15 | 225175_s_at | —1.4 | —0.4 |
| cg26143719 | C1QTNF6 | C1q and tumor necrosis factor-related protein 6 | No | 0.16 | 242444_at | 3.2 | 0.8 |
| cg04106785 | CDK5R1 | Cyclin-dependent kinase 5; regulatory subunit 1 | Yes | 0.18 | 204995_at | 1.4 | 0.6 |
| cg12120741 | EDNRB | Endothelin receptor type B | Yes | 0.21 | 202473_at | 2.2 | 0.4 |

**DISCUSSION**

**DNA Methylome**

Methylation of the carbon-5 position of cytosine, mostly in the context of CpG dinucleotides, is the main epigenetic modification of DNA [11, 13] and is essential for many biological processes [11, 13]. About 10%–20% of genes display DNA methylation patterns in a tissue-specific manner, and these are usually associated with tissue-specific patterns of gene expression [12]. Herein, in severe endometriosis, similar to other noncancerous somatic tissues including normal endometrium [14], most of the CGIs across the genome were hypomethylated, while a small group of CpG sites was hypermethylated, with most of them at non-CGIs. Of interest, most differential DNA methylation across the cycle in women with disease and in the comparison of disease versus controls involved CGIs, whereas changes in controls across the cycle...
involved both CGIs and non-CGIs, equally [14]. This is significant because CGI methylation is a highly controlled event in somatic tissues, and aberrant CGI methylation at key loci can alter cellular and tissue functions (e.g., tumor suppressor genes in cancer) [13, 32]. Moreover, various chromatin-remodeling components are involved in CGI methylation, and overrepresentation of abnormally methylated CGIs in endometriosis suggests wide-range abnormalities of the chromatin remodeling machinery in endometriosis.

Changes in DNA methylation across the cycle could result from preexisting cis or trans epigenetic differences that, together with vast hormone-induced transcriptome changes, could further alter the methylation status of a given CpG site. Of interest with regard to different hormonal milieu were the findings of the greatest differences between MSE<sub>Endo</sub> versus MSE<sub>Control</sub> (peak P<sub>4</sub>; Fig. 6a), when resistance to P<sub>4</sub> action is a hallmark of endometrial dysfunction in endometriosis [8]. Also, the greatest differences in disease were within the secretory phase (MSE<sub>Endo</sub> vs. ESE<sub>Endo</sub>; Fig. 6b, Endometriosis). These are different compared with controls, with greatest differences between peak P<sub>4</sub> versus peak E<sub>2</sub> (MSE<sub>Control</sub> vs. PE<sub>Control</sub>) and fewest differences in the secretory phase (MSE<sub>Control</sub> vs. ESE<sub>Control</sub>; Fig. 6b, Control), underscoring aberrant P<sub>4</sub> response in endometrium of women with endometriosis (Fig. 6 summarizes these findings and in relation to hormonal and endometrial changes across the menstrual cycle).

Although we found differences between phases and across the cycle in disease versus controls, the great majority of the platform’s CpG sites showed little to no difference from normal, and the differences were not very robust. Furthermore, the number of loci differentially methylated between all disease samples versus all control samples (phase independent, disease dependent) was small, similar to observations recently reported: Saare et al. [33], using endometrial tissue samples from patients with and without endometriosis and a larger DNA methylation platform, also found a small number of differentially methylated regions (DMRs) between disease and nondisease, and also with nonrobust DNA methylation changes. Yamagata et al. [34], using the same DNA

FIG. 4. CpG sites differentially methylated in eutopic endometrium across the cycle in disease. Columns represent samples. Rows represent loci. Green to red represents low to high methylation. Largest differences between the phases of cycle in disease are observed between MSE versus ESE, with fewer changes between MSE and PE and between ESE and PE. Top portion of each heat map represents loss of methylation in the earlier phases versus the later phase depicted in each comparison heat map, and the bottom portion represents gain of methylation. The location of each locus within or outside CGIs is depicted in the black (within) and white (outside) columns near each heat map. Loci in each comparison heat map are unique. Endo, endometriosis.
methylome platform used herein, observed no separation of genome-wide profiling of cultured endometrial stromal cells from patients with and without endometriosis in cluster or principal component analyses. In our study as well as that of Saare et al. using whole-tissue samples, this could be due to the variable DNA methylation levels across the cycle, or in different cellular components in endometrium, in both disease and controls. Although a small number of disease-specific differentially methylated loci, with nonrobust DNA methylation changes, limits marker discovery potentials, abnormal methylation patterns between phases of disease and across the cycle provide molecular insights for understanding the disease pathophysiology, as discussed below. It is important to note that further functional and protein expression analyses, following this exploratory study, are required to fully assess the role of epigenetic mechanisms in regulation of gene expression in these loci, as well as their potential importance in disease pathogenesis and progression.

**Insights into Endometriosis Pathophysiology**

Endometriosis is an inflammatory, estrogen-dependent disorder characterized by the presence of endometrial tissue outside the uterine cavity. It is accepted that endometrial cells by retrograde menstruation establish lesions on the peritoneum and escape immune clearance, with attachment, invasion, proliferation, and neoangiogenesis necessary for continued growth and survival of the endometriotic implants. These cellular advantages may be preexisting (hereditary or acquired) in eutopic endometrial cells of endometriosis patients. This is supported by the observation that eutopic endometrium from women with endometriosis differs in transcriptome, proteome, cellular signaling, and biochemical pathways [5, 8, 35] across the menstrual cycle, and shows a persistent proliferative and proinflammatory phenotype, increased cell survival, and aberrant P4 responsiveness in humans as well as in animal models of the disease [8, 10, 36, 37]. These studies have found specific pathways, gene networks, or candidate genes to be affected in disease that are potentially advantageous in disease establishment. Herein, aberrant DNA methylation profiles in eutopic endometrium affect genes with various functions potentially important in the pathogenesis and pathophysiology of the disease—such as cell cycle regulation; inflammation and immune response; steroid hormone response; cell migration; and regulation of gene expression. These include either identical or closely related genes reported in previous transcriptomic studies [8, 24, 27].

**Cell cycle, proliferation.** Our previous transcriptome studies in the eutopic endometrium of women with endometriosis versus no disease showed an enrichment of genes involved in cell cycle control and proliferation that are normally up-regulated in the proliferative phase and down-regulated in the secretory phase [8, 10]. These data suggested incomplete transition from the proliferative to the secretory phase, with enhanced cellular survival and aberrant expression of P4-regulated genes in secretory endometrium [8, 10]. Herein, we also observed abnormal DNA methylation associated with expression of genes regulating the cell cycle. For example, CDKN2B encodes a cyclin-dependent kinase inhibitor that binds specifically to CDK4 and CDK6 [38], both of which are abnormally expressed in endometrial stromal fibroblasts (eSFs) from women with endometriosis [25], and regulates cell proliferation by controlling cell cycle G1 progression [38]. The region encompassing CDKN2B and CDKN2A is affected by many signaling pathways and oncogenic agents [39], and it shows frequent copy number alterations, inactivating mutations, homozygous deletions, and aberrant DNA methylation/polycomb-associated silencing in a variety of tumors [38]. CDKN2B was more methylated in ESE compared with MSE in endometriosis and was associated with decreased expression, potentially leading to cell cycle abnormalities in ESE in disease.

Growth factors induce proliferation in steroid-dependent cells of normal endometrium. For example, the EGFR pathway is involved in eutopic and ectopic endometrial growth and differentiation, and eSF proliferation in vitro [8]. EGF is a potent mitogen for eSF in vitro [40], and MIG6, a negative regulator of EGFR signaling, is down-regulated in ESE in the endometrium of women with endometriosis [8]. Also, dysregulation of several genes in the EGFR signaling cascade was previously shown in patients [8] and in the baboon model of endometriosis [27]. Herein, we observed aberrant methylation of several EGF-like factors (TEK and LAMA3) in disease, further supporting disrupted EGF signaling in eutopic endo-
### ABERRANT ENDOMETRIAL EPIGENOME IN ENDOMETRIOSIS

#### TABLE 5. Select differentially methylated loci associated with gene expression across the phases in disease.

| Illumina probe ID | Gene symbol | Product | CGI | Median \( \beta \) value difference | Affymetrix probe set ID | Fold change | Spearman rho |
|-------------------|-------------|---------|-----|-------------------------------------|------------------------|-------------|--------------|
| **Endo ESE versus Endo PE** | | | | | | | |
| cg21561173 | C21orf81 | Hypothetical protein LOC114035 | Yes | 0.18 | 1569607_s_at | 3.7 | 0.6 |
| cg26267310 | HS317B14 | Dehydrogenase/reductase (SDR family) member 10 | Yes | 0.16 | 228713_s_at | -3.6 | -0.6 |
| cg19346899 | C16orf75 | Hypothetical protein LOC116028 | Yes | -0.14 | 226456_at | -1.6 | 0.7 |
| cg19353083 | RAD50 | RAD50 homolog isoform 1 | Yes | -0.14 | 209349_at | 2.1 | -0.3 |
| cg02248486 | HOXA5 | Homeobox A5 | Yes | -0.18 | 213844_at | 2.2 | -0.8 |
| cg09324116 | GEMIN8 | Family with sequence similarity 51; member A1 | Yes | -0.24 | 222854_s_at | 1.9 | -0.4 |
| **Endo ESE versus Endo MSE** | | | | | | | |
| cg25608041 | TBC1D1 | TBC1 (tre-2/USp6; BUB2; cdc16) domain family member 1 | Yes | -0.18 | 212350_at | -1.6 | 0.6 |
| cg17826679 | SLC4A42 | CTL2 protein | Yes | -0.18 | 221753_s_at | -1.4 | 0.5 |
| cg15784615 | LTBR | Lymphotixin beta receptor | No | -0.16 | 203005_at | -2.2 | 0.8 |
| cg121020741 | EDNRB | Endothelin receptor type B isoform 2 | Yes | -0.16 | 204273_at | -24.0 | 0.7 |
| cg121020741 | EDNRB | Endothelin receptor type B isoform 2 | Yes | -0.16 | 206701_s_at | -18.2 | 0.5 |
| cg06339706 | PLEKHA4 | Pleckstrin homology domain containing family A (phosphoinositide-binding-specific) member 4 | No | -0.16 | 219011_at | -1.9 | 0.6 |
| cg09835085 | KCNE4 | Potassium voltage-gated channel; Isk-related family; member 4 | Yes | -0.16 | 222379_at | 1.8 | -0.9 |
| cg22038738 | PLAT | Plasminogen activator; tissue type isoform 1 preproprotein | No | -0.15 | 201860_s_at | -3.7 | 0.9 |
| cg26780404 | COL12A1 | Alpha 1 type XII collagen short isoform precursor | Yes | -0.15 | 225664_at | -3.2 | 0.4 |
| cg26780404 | COL12A1 | Alpha 1 type XII collagen short isoform precursor | Yes | -0.15 | 233109_at | -2.0 | 0.8 |
| cg25527547 | PLOD3 | Procollagen-lysine-2-oxoglutarate 5-dioxigenase 3 precursor | No | -0.15 | 202185_at | -1.8 | 0.8 |
| cg12969602 | TAF1D | TATA-box-binding protein associated factor, RNA polymerase I, subunit D | No | -0.15 | 221580_s_at | 2.8 | -0.9 |
| cg09324116 | GEMIN8 | Family with sequence similarity 51; member A1 | Yes | -0.14 | 219252_s_at | 2.2 | -0.5 |
| cg00131557 | DNAJC15 | DNAJ domain-containing | Yes | -0.14 | 218435_at | 2.9 | -0.4 |
| cg25280512 | NCO2 | Nuclear receptor corepressor 2 | Yes | -0.14 | 207760_s_at | -1.3 | 0.9 |
| cg00995152 | DAB2IP | DAB2-interacting protein isoform 2 | Yes | -0.14 | 225020_at | -2.0 | 0.6 |
| cg26143719 | CIQTNF6 | CIQ and tumor necrosis factor-related protein 6 | No | -0.14 | 223571_at | -11.6 | 0.9 |
| cg26143719 | CIQTNF6 | CIQ and tumor necrosis factor-related protein 6 | No | -0.14 | 242444_at | -8.2 | 0.9 |
| cg17740645 | GRB7 | Growth factor receptor-bound protein 7 | No | -0.14 | 210761_s_at | -8.5 | 0.7 |
| cg19481866 | CKDKN2B | Cyclin-dependent kinase inhibitor 2B isoform 2 | Yes | 0.14 | 236313_at | -5.2 | -0.6 |
| cg08097657 | SEMA3B | Semaphorin 3B isoform 2 precursor | Yes | 0.14 | 203071_at | -3.2 | -0.9 |
| cg16529592 | RUNX3 | Runt-related transcription factor 3 isoform 2 | Yes | 0.14 | 204198_s_at | -10.5 | -0.7 |
| cg02794695 | SLA | Srr-like adaptor | No | 0.15 | 203761_at | 6.8 | -0.9 |
| cg17518965 | S1PR4 | Endothelial differentiation; G protein-coupled receptor 6 precursor | Yes | 0.15 | 206437_at | -2.2 | -0.5 |
| cg22518733 | CCL3 | Chemokine (C-C motif) ligand 3 | No | 0.15 | 205114_s_at | -5.2 | -0.8 |
| cg06132676 | AIF3 | Apoptosis-inducing factor family; member 3 isoform 1 precursor | No | 0.15 | 227919_at | 9.4 | 0.9 |
| cg121561173 | C21orf81 | Hypothetical protein LOC114035 | Yes | 0.16 | 1369607_s_at | 1.6 | 0.8 |
| cg17657618 | RP1-21O18.1 | Hypothetical protein LOC39563 | Yes | 0.16 | 213478_at | 4.3 | -0.7 |
| cg04872689 | PLEK | Pleckstrin | No | 0.18 | 203471_s_at | -2.4 | -0.8 |
| cg16363586 | BST2 | Bone marrow stromal cell antigen 2 | Yes | 0.18 | 203471_at | 6.7 | -0.8 |
| **Endo MSE vs. Endo PE** | | | | | | | |
| cg27627350 | ALDH1A3 | Aldehyde dehydrogenase 1A3 | Yes | -0.22 | 203180_at | 18.8 | -0.4 |
| cg18236477 | ATP8A2 | ATPase, aminophospholipid transporter-like; Class I; type 8A; member 2 | Yes | -0.20 | 231395_at | 2.3 | -0.3 |
| cg25228126 | FZD2 | Frizzled 2 | No | -0.19 | 201220_at | -2.5 | 0.3 |
| cg09835085 | KCNE4 | Potassium voltage-gated channel; Isk-related family; member 4 | Yes | 0.16 | 222379_at | -1.7 | -0.5 |
| cg04872689 | PLEK | Pleckstrin | No | -0.15 | 203471_s_at | 2.8 | -0.5 |
| cg13673094 | PTPRC | Protein tyrosine phosphatase; receptor type C isoform 1 precursor | No | 0.15 | 212588_at | 5.4 | -0.6 |
| cg22678136 | SNRPN | Small nuclear ribonucleoprotein polypeptide N | Yes | -0.14 | 206042_x_at | -1.6 | 0.5 |
| cg01009664 | TRH | Thyrotropin-releasing hormone | Yes | -0.18 | 206622_at | -49.0 | 0.7 |
| cg02184413 | VNN1 | Vanin 1 precursor | No | -0.16 | 205844_at | 6.9 | -0.4 |
metrium of women with endometriosis. TEK (tyrosine kinase, endothelial) plays a role in angiogenesis [41] and is more methylated in PE in disease versus controls, with decreased expression. LAMA3 (Laminin, alpha 3), an extracellular matrix component of basement membranes promoting epithelial cell attachment, migration, and organization [42], is less methylated and more expressed in MSE in disease versus control. Inflammation and immune response. Cytokines and growth factors effectively promote endometrial cell/tissue fragment implantation and evasion of immune-mediated clearance, playing a pivotal role in endometriosis pathophysiology [40]. Activated macrophages are more abundant, and cytokine/chemokine profiles are altered in peritoneal fluid and eutopic endometrium of endometriosis patients [43, 44] and in baboons with disease, including several aberrantly expressed

FIG. 6. Summary figure. Top portion is a graphic of changes in estrogen, progesterone, and endometrial characteristics across the menstrual cycle depicting the PE, ESE, and MSE phases. a) Summary of phase-specific changes in disease versus control, showing the largest differences between MSE, compared with PE and ESE, with more gain of methylation loci in disease compared with control in PE, and more loss of methylation in disease compared with control in ESE and MSE. b) Comparison of changes in DNA methylation across the cycle and the comparison with control, showing the largest difference in disease between MSE versus ESE different from control, with smallest differences between MSE versus ESE and largest differences in MSE versus PE. Also, changes involved loss and gain of methylation equally in control in all three comparisons, different from disease showing secretory phases (ESE and MSE) involving more loss of methylation compared with PE and more gain of methylation in MSE versus ESE. Endo, endometriosis.
Genes involved in T-cell activation [27]. Transcriptomic studies demonstrate a proinflammatory phenotype in women with endometriosis versus controls without any pelvic pathology [10]. Herein, several differentially methylated genes have roles in inflammation and the immune response, including BST2, PLEK, and CCL3. BST2 (bone marrow stromal cell antigen 2, also known as CD317) is an interferon-induced protein. Inflammatory cytokines, such as interleukin 6 (IL6), induce its expression, and it is highly overexpressed in endometrial cancer [45]. It is less methylated and more expressed in MSE\textsubscript{Endo} than ESE\textsubscript{Endo}. PLEK (pleckstrin) plays an important role in proinflammatory cytokine activation (tumor necrosis factor \(\alpha\) and IL1B) [46]. Its overexpression alters cytoskeletal organization and cell spreading [47]. It is less methylated and more expressed in MSE\textsubscript{Endo} versus ESE\textsubscript{Endo} and PE\textsubscript{Endo}, and also less methylated and more expressed in MSE\textsubscript{Endo} than MSE\textsubscript{Control}. CCL3, chemokine (C-C motif) ligand 3 (also known as macrophage inflammatory protein 1\(\alpha\) [MIP1\(\alpha\)]) is elevated in the peritoneal fluid of women with endometriosis [48]. In vitro studies suggest its role in inducing monocyte and T-cell migration to ectopic sites and increasing eSF invasiveness and endometriosis progression through metalloproteinase induction [48]. High E\(_2\) levels (in endometriosis) further induce its expression [48]. It is less methylated and highly upregulated in MSE\textsubscript{Endo} versus ESE\textsubscript{Endo}.

Other gene groups potentially important in disease pathophysiology and/or progression. Differentially methylated genes affecting diverse processes at the cellular and molecular levels include those involved in cell adhesion (CDH8, CCDN8, COL12A1, and COL11A2), cell migration (G1PC1, CCKAR, and TN3), responses to hormones (PRSS8 and TFF3), protein kinase cascades (IKBKEX and SLC44A2), and oxidative stress protection (ALDH1A3, CYP2A13, GSTM1, and PLOD3). Furthermore, many differentially methylated genes are involved in transcription function and regulation, and DNA binding and cation binding.

Previously Reported Genes with Abnormal Methylation in Disease

Several genes, including SF-1, ER-\(\beta\), HOXA10, and PR-B, are differentially methylated in the endometrium of women with endometriosis versus controls [15–17]. CpG sites interrogated in the current platform did not show the same patterns for these genes, because CpG methylation results differ based on the interrogated location. Also, every CpG within a region may not be methylated or may show variable methylation. aberrant methylation was found when interrogating 29 CpG sites across a 333-bp region containing exon II, intron II, exon III, and intron III (+4085 to +4337) of the SF-1 gene in endometriosis [17]. However, our platform interrogated two locations (at +17 and +127), and both probes are unmethylated in disease and control. In another study, when HOXA10 methylation was interrogated at three regions in disease (−25 to −300 and other fragments within 1.2 kb of the transcription start site), it showed no, partial, or full methylation [16]. Our platform has one probe for HOXA10 at +1461, a different location, and shows intermediate methylation in controls and disease, and no differential methylation, thus accounting for the observed differences. Naqvi et al. [49] used the Illumina platform and reported on 10 of 120 genes differentially methylated, and none of these 10 genes was in common with our data, potentially because of the use of samples from individuals with different severities of disease and/or different cycle phases, or, as they recently found, the distance of disease from the uterus [50]. Saare et al. [33], as mentioned above, have shown similar genome-wide profiles between eutopic endometrium of disease and nondisease as those reported herein. They found the largest differences in the menstrual and late secretory phases (not part of the current study) versus other cycle phases, and although they showed there are changes in methylation across the cycle, they did not find many differences between the proliferative versus the early secretory and midsecretory phases. There are several study differences that could affect the specific outcomes of our study compared with Saare et al. [33], including in their report: a smaller number of control samples in the proliferative and early secretory phases; defining and restricting the search for differential methylation to DMRs (regions smaller than 500 bp between at least three consecutive probes); a larger array platform; self-reporting of cycle date for most samples, which may affect phase determination accuracy; and the availability of samples in the menstrual and late secretory phases.

Strengths and Limitations of the Study

Strengths of our study include well-annotated participant data; strict cycle phase determination using histology, serum \(E_2\) and \(P_4\) levels, and unsupervised hierarchical clustering and principal component analyses; focus on the most severe phenotype of disease by ASRM criteria [19]; and concomitant gene expression and DNA methylation analysis on each endometrial sample from the same participants. Limitations include the small sample size, cellular heterogeneity, and the cross-sectional study design. However, despite these caveats, we observed the biggest differences in the DNA methylome in the secretory phase—the midsecretory phase showing the largest differences when comparing disease to controls and in midsecretory versus early secretory phases across the cycle in disease—in line with abnormal \(P_4\) response in endometriosis. In addition, we found differentially methylated genes associated with gene expression changes relevant to the pathophysiology of endometriosis. However, in the absence of functional analyses, our interpretation of the impact of our findings of DNA methylation differences in both gene expression regulation and disease pathogenesis is limited. Cellular heterogeneity due to the use of whole-tissue samples adds to potential variations in DNA methylation values, meaning that the data are drawn from a wide range of probability distributions (different cell types predicted to have different methylation profiles), which is particularly important with a small sample size and not normally distributed data. We removed samples from further analyses only if they failed our stringent, multistep QC’s, and we used a more resilient statistical method to potential outliers that was also suitable for nonparametric data, but we did not cherry-pick the samples for a more homogenous signature. A larger sample set, collected longitudinally, with isolated cellular components and using a comprehensive genome-wide DNA methylation and gene expression platform in future studies, together with protein expression and functional analyses, will extend the understanding of epigenetic abnormalities in the eutopic endometrium of women with endometriosis.

REFERENCES

1. Giudice LC. Clinical practice: endometriosis. N Engl J Med 2010; 362: 2389–2398.
2. Bulun SE. Endometriosis. N Engl J Med 2009; 360:268–279.
3. Guo SW. Epigenetics of endometriosis. Mol Hum Reprod 2009; 15: 587–607.
4. Arimoto T, Katagiri T, Oda K, Tsunoda T, Yasugi T, Osuga Y, Yoshikawa H, Nishii O, Yano T, Taketani Y, Nakamura Y. Genome-wide cDNA
microarray analysis of gene-expression profiles involved in ovarian endometriosis. Int J Oncol 2003; 22:551–560.
5. Kao LC, Germeyer A, Tulac S, Lobo S, Yang JP, Taylor RN, Osteen K, Levine LS. Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility. Endocrinology 2003; 144:2870–2881.
6. Matsuzaki S, Canis M, Vaurs-Barriere C, Boesvlag-Tanguy O, Dastugue B, Mage G. DNA microarray analysis of gene expression in eutopic endometrium from women with deep endometriosis using laser capture microdissection. Fertil Steril 2005; 84:suppl 2:1180–1190.
7. Van Langendonckt A, Punyadeera C, Kamps R, Dunselman G, Klein-Hitpass L, Schurgers LJ, Squifflf J, Donnez J, Groothesis P. Identification of novel antigens in blood vessels in rectovaginal endometriosis. Mol Hum Reprod 2007; 13:875–886.
8. Burney RO, Talbi S, Hamilton AE, Vo KC, Nyegaard M, Nezhat CR, Lessey BA, Giudice LC. Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis. Endocrinology 2007; 148:3814–3826.
9. Zafarakas M, Tarlatzis BC, Streichert T, Pournaropoulos F, Wolfe U, Smets SJ, Wittek B, Grimbizis G, Braakhoff RN, Panteli K, Bonits J, Gunes C. Genome-wide microarray gene expression, array-CGH analysis, and telomerase activity in advanced ovarian endometriosis: a high degree of differentiation rather than malignant potential. Int J Mol Med 2008; 21:335–344.
10. Tamaresis JS, Irwin JC, Goldfien GA, Rabban JT, Burney RO, Nezhat C, Sun J, Olianas A, Nezhat CN, Conti P, et al. Genomic and transcriptomic changes in eutopic endometrium of patients with endometriosis. Am J Obstet Gynecol 2005; 193:371–380.
11. Kossar S, Zelenko Z, Irwin JC, Giudice LC. DNA methylation patterns of steroid receptor genes ESR1, ESR2 and PGR expression regulation. Mol Endocrinol 2014; 28:1118–1135.
12. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. Nat Rev Genet 2012; 13:484–492.
13. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012; 13:484–492.
14. Houshdaran S, Zelenko Z, Irwin JC, Giudice LC. DNA methylome is cycle-dependent and is associated with gene expression regulation. Mol Endocrinol 2014; 28:1118–1135.
15. Meyer JL, Zimbardi D, Podgaec S, Amorim RL, Abrao MS, Rainho CA. Splicing of the Dec1 transcript in normal endometrium and eutopic endometriosis. Fertil Steril 2009; 150:1461–1466.
16. Popov N, Gil J. Epigenetic regulation of the INK4b-ARF-INK4a locus: in sickness and in health. Epigenetics 2010; 5:685–690.
17. May KE, Villar J, Kirtley S, Kennedy SH, Becker CM. Endometrial alterations in endometriosis: a systematic review of putative biomarkers. Hum Reprod Update 2011; 17:637–653.
18. Brooks J, Mangdarne AG, Fazleabas AT. The non-human primate model of endometriosis: research and implications for fecundity. Mol Hum Reprod 2010; 15:577–586.
19. Fazleabas AT. Progesterone resistance in a baboon model of endometriosis. Semin Reprod Med 2010; 28:75–80.
20. Margari KM, Zafiropoulos A, Hatzidaki E, Giannakopoulou C, Arici A, Mavropoulou A, Manifavas C. Changes in eutopic endometrial gene expression during the progression of endometriosis. Biol Reprod 2011; 85:1003–1011.
21. Aghajanova L, Horcajadas JA, Weeks JL, Esteban FJ, Nezhat CN, Conti M, Giudice LC. The protein kinase A pathway-regulated transcriptome of endometrial stromal fibroblasts reveals compromised differentiation and persistent proliferative potential in endometriosis. Endocrinology 2010; 151:1341–1355.
22. Aghajanova L, Horcajadas JA, Weeks JL, Esteban FJ, Nezhat CN, Conti M, Giudice LC. The protein kinase A pathway-regulated transcriptome of endometrial stromal fibroblasts reveals compromised differentiation and persistent proliferative potential in endometriosis. Endocrinology 2010; 151:1341–1355.