Endometriosis Gene Expression Heterogeneity and Biosignature: A Phylogenetic Analysis

Mones Abu-Asab,1 Ming Zhang,2 Dennis Amini,3 Nihad Abu-Asab,4 and Hakima Amri2

1 Laboratory of Immunology, Section of Immunopathology, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA
2 Department of Biochemistry and Cellular & Molecular Biology, Georgetown University Medical Center, Washington, DC 20007, USA
3 Department of Obstetrics and Gynecology, Georgetown University Hospital, Washington, DC 20007, USA
4 Armidale Rural Referral Hospital, University of New England and The University of Newcastle, Armidale, NSW 2351, Australia

Correspondence should be addressed to Hakima Amri, amrih@georgetown.edu

Received 13 July 2011; Accepted 13 September 2011

Academic Editor: Kimberly K. Leslie

Copyright © 2011 Mones Abu-Asab et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Endometriosis is a multifactorial disease with poorly understood etiology, and reflecting an evolutionary nature where genetic alterations accumulate throughout pathogenesis. Our objective was to characterize the heterogeneous pathological process using parsimony phylogenetics. Gene expression microarray data of ovarian endometriosis obtained from NCBI database were polarized and coded into derived (abnormal) and ancestral (normal) states. Such alterations are referred to as synapomorphies in a phylogenetic sense (or biomarkers). Subsequent gene linkage was modeled by Genomatix BiblioSphere Pathway software. A list of clonally shared derived (abnormal) expressions revealed the pattern of heterogeneity among specimens. In addition, it has identified disruptions within the major regulatory pathways including those involved in cell proliferation, steroidogenesis, angiogenesis, cytoskeletal organization and integrity, and tumorigenesis, as well as cell adhesion and migration. Furthermore, the analysis supported the potential central involvement of ESR2 in the initiation of endometriosis. The pathogenesis mapping showed that eutopic and ectopic lesions have different molecular biosignatures.

1. Introduction

In the USA, 10–20% of women suffer from endometriosis, with 40% developing infertility. It is a public health issue with a patient’s medical costs approximately 63% higher than those of the average woman [1]. The etiology and pathophysiology of endometriosis remains poorly understood. The hypothesis of retrograde menstruation is the oldest and most widely accepted. However, it fails to explain why some women develop endometriosis while others do not, given that some degree of retrograde menstrual flow is experienced by all women [2]. Thus, other factors such as immunology [3, 4], genetics [5], and the environment [6] may play a role in the mechanism of disease development.

The genetic theory dates back to the first formal genetic study published in 1980 by Simpson et al. [5]. Numerous findings since then support a polygenic multifactorial inheritance of endometriosis caused by an interaction between multiple genes with the environment. Several studies from the OXGENE (Oxford Endometriosis GENE) group confirmed an inheritance component to endometriosis. Specifically, in one report of 100 families with endometriosis from 6 different countries, 19 mother-daughter pairs and 56 sibling pairs shared the disease [7].

A diagnostic method that screened for genetic profiles or candidate genes may benefit the patient by detecting disease earlier, improving patient quality of life, discerning genetic predisposition, lowering costs, and reducing the need for invasive laparoscopic investigations.

Though not yet completely understood, numerous studies show a correlation between the occurrence of endometriosis and ovarian cancer [8–10]. Both diseases share pathogenic factors such as familial predisposition, genetic modifications, immunologic abnormalities, uncontrolled angiogenesis, and hormonal disturbances [11]. Malignant transformation of endometriosis has been reported [2, 12]. It
is hypothesized that ectopic glands may expand monoclonally; however, this phenomenon is not yet clearly defined [13]. Elucidation of a cellular continuum from benign endometriosis to malignancy requires more research and a greater understanding of common mutational events.

Molecular processes involved in disease development share aspects of evolutionary transformation such as genetic mutations, clonal propagation, irreversible gene expression, and shared derived genetic alterations. Sarnat and Netsky first put forth the concept of disease etiology by evolutionary criteria in 1984 [14] whereby disease is viewed as an accumulation of genetic mutations. In this study, we sought to identify a genomic biosignature(s) for endometriosis using a newly developed evolution-based parsimony phylogenetics approach for gene expression microarrays data [15, 16] of endometriosis patients in order to stratify individual cases based on the molecular change, model the disease based on the level of patients’ gene expression profiles, and identify affected molecular pathways involved in the disease process.

2. Methods

Gene expression microarray datasets of endometriosis patients, GSE7305, from NCBI’s Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) were used in the study [17]. Briefly, the datasets, submitted to NCBI by Hever and colleagues and successfully published in the Proc. Natl. Acad. Sci. USA [17], comprised 10 ovarian endometriosis and 10 matched eutopic endometria from the same patients using Affymetrix Human Genome U133 Plus 2.0 gene array. Polarity assessment was carried out using UNIPAL (Universal Parsing Algorithm) [18] by coding the expression values into ancestral (unchanged) and derived (deregulated/abnormal) states. Unchanged values were coded as zero (0) and altered/deregulated values as one (1), thus transforming the original expression values into a qualitative binary matrix of 0’s and 1’s [18]. Hierarchical classification through parsimony phylogenetic analysis was carried out with MIX, the maximum parsimony program of PHYLIP ver. 3.57c to produce cladograms [19]. TreeView was used to redraw the cladogram for final illustration [20].

Clonal alterations (or shared derived expression states, named synapomorphies in the field of phylogenetics or biomarkers in the biomedical field) were utilized to delimit a natural group of related specimens, termed a “clade.” The tree-like diagram, the cladogram, is the classification hypothesis that models the relatedness between the specimens. The full list of synapomorphies circumscribing the diseased specimens was extracted and analyzed without a priori selection. The analysis modeled the patterns of change occurring in the gene expression data set, classified specimens, and mapped the molecular events of altered pathways.

The synapomorphies were then modeled by Genomatix BiblioSphere Pathway Edition software version 7.2 for gene linkage. Genomatix BiblioSphere is data-mining software that extracts and analyzes gene relationships from literature databases (primarily NCBI PubMed) and genome-wide promoter analysis. The Genomatix collection of gene names and synonyms are supplied by NCBI Locust Link. We used this program to model our synapomorphies to reveal gene linkage and the affected pathways of pathogenesis. The gene maps have been filtered with respect to abstract cocitation level, B0.

3. Results

The maximum parsimony phylogenetic program, MIX, produced one most parsimonious cladogram (Figure 1(a)). The cladogram has a directionality that reveals the amount of accumulated shared derived (abnormal) expressions from the base up (from eutopic endometrium to ectopic endometriosis specimens). The endometriosis specimens have the highest number of abnormal gene expressions. The analysis revealed that all the endometriosis specimens share 3,636 synapomorphies (or shared derived (abnormal) gene expressions) that distinguish them from the eutopic endometrium specimens. Synapomorphies are additive and accumulate for specimens positioned higher along the main axis of the cladogram. Thus, the additional 1,923 synapomorphies characterizing the highest clad consisting of 4 specimens (GSM175766, GSM175767, GSM175769, and GSM175770) share the greatest amount of deregulated genetic information that is most specific to these four subjects. Within the directionality or continuum of change, eutopic endometrium specimens (GSM175783–85) that may be susceptible to developing into endometriosis are located between the largest eutopic endometrium clade (GSM175777–78) and first endometriosis clade (GSM175775). The cladogram identified transitional patterns from eutopic endometrium (Figure 1(a)—lower, green) to endometriosis (Figure 1(a)—upper, black) specimens (GSM175783–85); these specimens separated from the majority of eutopic endometrium specimens and formed a transitional zone closer to the beginning of endometriosis specimens’ clade. This analytical method can graphically delineate and molecularly represent the progression of endometriosis through the accumulation of changes in gene expression.

The modeling of gene expression heterogeneity is illustrated in the heat map in Figure 1(b). The heat map depicts overexpressed (red), underexpressed (yellow), and unchanged (green) gene expressions of 16 selected genes in all ten endometriosis specimens and their relative expression pattern per specimen, thus demonstrating the differential expression of genes across specimens.

The expression of ZC3H15, taken as an example (Figure 1(b), gene number 16), typifies the heterogeneity of single gene expression among the 10 endometriosis specimens. Thus, among the endometriosis patients, ZC3H15 has a heterogeneous expression: while it is unchanged in 4 specimens (GSM175766, GSM175774, GSM175768, and GSM175773), it is overexpressed in GSM175769 and underexpressed in 5 specimens (GSM175767, GSM175765, GSM175772, GSM175771, and GSM175775). Furthermore, the horizontal frame denotes the heterogeneity of several of the 16 genes highlighted in the heat map even within two specimens in close proximity on the cladogram that share a multitude of synapomorphies.
The robustness of parsimony phylogenetics to model genetic heterogeneity is further illustrated in Figure 2. Lipo- calin 2 (LCN2) and MYB binding protein (P160, MYBBP1A) are two examples of genes with dichotomous expression (above and below their gene expression range of normal specimens) as well as heterogeneity within the normal range of gene expression (Figure 2(a)). MYBBP1A appears to be tightly regulated as a slight deviation from the normal range which seems to induce a pathological state (Figure 2(b)).

The functional and regulation relationships of the differentially expressed genes were assessed using Genomatix BiblioSphere (http://www.genomatix.com/). This analysis focused on the 1,923 synapomorphic genes of the highest clade (GSM175766, GSM175767, GSM175769, and GSM175770) to yield the greatest wealth of genomic insight into the pathology of endometriosis. Groups of shared derived genes were entered into Genomatix BiblioSphere including underexpressed and overexpressed. We analyzed the gene maps for prominent nodes as well as their central and extended linkages.

Out of the 1,923 gene synapomorphies aforementioned, 583 overexpressed, coded in red (Figure 3(a) and full gene listing in Supplemental Material (1) see Supplementary Materials available on line at doi: 10.1155/2011/719059) and 459 underexpressed genes, coded in yellow (Figure 3(b) and full gene listing in Supplemental Material (2)) were modeled.

The cladogram in Figure 1 shows a group of three in- tandem specimens (GSM175783–85) that forms a transitional zone between eutopic endometrium and ectopic endometriosis. This clade was circumscribed by 707 synapomorphies. The pathway network analysis pointed out to the overexpressed ERS2 as the central deregulated gene affecting other gene nodes (Figure 4). This pathway analysis showed that the gene network was also different from the lower clade of the eutopic endometrial specimens (GSM175776, GSM175777, GSM175778, GSM175779, GSM175780, GSM175781, and GSM175782) (Figure 5).

4. Discussion

Parsimony phylogenetics, an evolution-based bioinformatic paradigm, revealed deregulated clonal expressions within ectopic endometriosis as compared to eutopic endometrium specimens. This analytical method achieved several goals: construction of the molecular disease boundaries and
Figure 2: Parsimony phylogenetics identifies expression heterogeneity of single genes. (a) LCN2 (lipocalin 2) and (b) MYBBP1A (MYB binding protein (P160)) depict the dichotomous (under and over) gene expression as well as heterogeneity within the range of gene expression of the specimens.

Figure 3: Gene linkage map for the over- (a) and underexpressed (b) genes filtered separately at the B0 level using Genomatix BiblioSphere. Gene list was obtained from the 4 specimens (GSM175766, GSM175767, GSM175769, and GSM175770) at the upper end of cladogram in Figure 1(a).

pathways’ aberrations, stratification (subtyping) of disease, detection of early disease stages, suggestion of potential therapeutic targets, and identification of the biosignature (profile) of diseased specimens.

The comprehensive parsimony phylogenetics analysis revealed an extensive list of shared derived (deregulated/abnormal) expression states—or synapomorphies in a phylogenetic sense (biomarkers in a biomedical sense), which showed the extent of heterogeneity among specimens. Furthermore, it identified dichotomously expressed asynchronous genes (DEA) among endometriosis specimens [16]; these are gene expression values that are above and below the range of gene expression of the eutopic endometrium specimens (Figure 1(b)). Each DEA gene splits the specimens into two groups, thus showing the heterogeneity that exists among endometriosis specimens. This pattern was first reported by Lyons-Weiler et al. [21] and termed DEA by Abu-Asab et al. This phenomenon was designated dichotomously expressed asynchronicity to reflect its two-tailed distribution and deviation from the expression range of the outgroup [16]. While F- and t-statistics as well as fold change may not consider DEA genes significant or include them within the list of differentially expressed genes [21], the polarity assessment algorithm assesses each value as
either derived or ancestral, thereby revealing the gene’s status in relation to the gene profile of the outgroup [16].

LCN2 and MYBBP1A heterogeneous expression as DEA genes illustrates the complexity of this disease. LCN2 is known as a marker from benign to pre- and malignant ovarian tumors and may be involved in progression of epithelial ovarian malignancies. It is also an epithelial inducer in Ras malignancies and a suppressor of metastasis [22]. Upregulated in ovarian cancer cells, it may be involved in the progression of epithelial ovarian malignancies [23]. Our results showed that 4 specimens exhibited LNC2 overexpression which could explain the risk of progression of endometriosis from a benign to malignant condition in some patients [13, 24].

MYBBP1A is a novel NF-kappaB corepressor of transcription and DNA-directed polymerase activity [25]. Associations between the p160 coactivator proteins and endocrine resistance have been described, involving the MAP kinase effector proteins Ets [26]. This corepressor gene expression appears to be tightly regulated as a slight deviation from the normal range appears to induce a pathological state.

4.1. Overexpressed Genes. From the list of overexpressed genes, we selected to discuss only a few among those reported in the literature as relevant to the pathogenesis of endometriosis.

The endocrine-gland-derived vascular endothelial growth factor (PROK1) has been shown to possess a paracrine role for prokinetics and their receptors in endometrial vascular function [27]. Endometriotic implants require neovascularization to proliferate and invade into ectopic sites, and such angiogenic factors are currently being targeted for novel medical therapeutics [28].

Caveolin-1 (CAV1) has been shown to negatively regulate the Jak-2 tyrosine kinase in mice [29] and the latter modulating the processes of cell proliferation, differentiation, and apoptosis [30].

Nerve growth factor (NGF) levels are higher in the follicular fluid of women with endometriosis [31]. Histological analysis of human deep innervating endometriosis (DIE) tissue showed strong expression of NGF in endometriotic glands and stroma of DIE which may play a role in the pathways involved in the intense pelvic pain that patients experience [32].

Hydroxysteroid (17-beta) dehydrogenase 11 (HSD17β-11) converts 5 alpha-androstane-3 alpha, 17 beta-diol to androsterone [33]. Expression analysis has revealed significant upregulation of enzymes involved in estradiol synthesis (i.e., aromatase, sulfatase, and all reductive 17 beta-HSDs), which indicates increased local levels of mitogenic estradiol and decreased levels of protective progesterone in endometriosis [34].

4.2. Underexpressed Genes. BCL-2-related ovarian killer (BOK) is a proapoptotic protein identified in the ovary [35] and functions as an essential mediator of p53-dependent apoptosis [36].

It is well established that the matrix metalloproteinase system (MMPs) plays an important role in the normal...
Table 1: Gene list of over- and underexpressed genes summarized by function. These dichotomously expressed genes reflect the gene expression heterogeneity among specimens.

| Function                      | Gene symbol |
|-------------------------------|-------------|
| **Cell cycle**                |             |
|                               | CAV1        |
|                               | GPC6        |
|                               | NCAP1       |
|                               | SLBP        |
|                               | CAV1        |
|                               | CCNA2       |
|                               | DDR2        |
|                               | IFI16       |
|                               | MAP2K1      |
|                               | MAD2L2      |
|                               | MAPK6       |
|                               | GPC3        |
|                               | GPC6        |
|                               | GTSE1       |
|                               | PTPN21      |
|                               | PTPN3       |
|                               | RBBP8       |
|                               | RPS6KA3     |
|                               | SLBP        |
|                               | TGFBR3      |
|                               | TP53        |
|                               | TRMU        |
| **Cell adhesion and migration** |             |
|                               | CDH1        |
|                               | CDH3        |
|                               | HOX-A10     |
|                               | HOX-A11     |
|                               | HOX-B5      |
|                               | HOX-C5      |
|                               | BOC         |
|                               | ICAM1       |
|                               | RIPK2       |
|                               | CDH2        |
|                               | CDH3        |
|                               | CLDN3       |
|                               | CLDN4       |
|                               | CLEC10A     |
|                               | HOXB2       |
|                               | HOXB3       |
|                               | HOXB4       |
|                               | HOXB7       |
|                               | HOXB8       |
|                               | HOXC4       |
|                               | ACTA2       |
|                               | ACTG2       |
|                               | ACTG1       |
|                               | CNKR1       |
|                               | DCLK1       |
|                               | ITGA11      |
|                               | ITGB8       |
|                               | LAMC2       |
|                               | PAK6        |
|                               | SPC25       |
|                               | BOC         |
|                               | ICAM1       |
|                               | RIPK2       |
|                               | CD40        |
|                               | JAK3        |
|                               | SIGLEC1     |
|                               | COL10A1     |
|                               | ESR1        |
|                               | KRT19       |
|                               | LAMA5       |
|                               | LAMC3       |
|                               | LAMB2       |
|                               | LAMC6       |
|                               | LTBP2       |
|                               | MMP26       |
|                               | NID2        |
|                               | PAK6        |
|                               | SPC25       |
|                               | TGFBR3      |
|                               | TNS1        |
|                               | VAPA        |
| **Apoptosis**                 |             |
|                               | IHPK2       |
|                               | BIRC5       |
|                               | BNIP3L      |
|                               | BOK         |
|                               | FAIM3       |
|                               | IL24        |
|                               | MAP2K6      |
|                               | MAPK10      |
|                               | PK6         |
|                               | RBMS3       |
|                               | TP53        |
|                               | TP52L1      |
|                               | WDR26       |
| **Infertility/fertility**     |             |
|                               | LEFTY2      |
|                               | PK6         |
|                               | SPA17       |
|                               | SPAG1       |
| **Immunity**                  |             |
|                               | BOC         |
|                               | ICAM1       |
|                               | RIPK2       |
|                               | BOC         |
|                               | BST2        |
|                               | CD40        |
|                               | CEACAM1     |
|                               | PVR1        |
|                               | IP39A1      |
|                               | SIGLEC1     |
|                               | COL10A1     |
|                               | ESR1        |
|                               | KRT19       |
|                               | LAMA5       |
|                               | LAMC3       |
|                               | LAMB2       |
|                               | LAMC6       |
|                               | LTBP2       |
|                               | MMP26       |
|                               | NID2        |
|                               | PAK6        |
|                               | SPC25       |
|                               | TGFBR3      |
|                               | TNS1        |
|                               | VAPA        |
| **Cell structure**            |             |
|                               | ACTA2       |
|                               | ACTG2       |
|                               | CNKR1       |
|                               | DCLK1       |
|                               | ITGA11      |
|                               | LAMC2       |
|                               | PAK6        |
|                               | SPC25       |
|                               | ACTA2       |
|                               | ACTG2       |
|                               | ACTG1       |
|                               | CNKR1       |
|                               | DCLK1       |
|                               | ITGA11      |
|                               | LAMC2       |
|                               | PAK6        |
|                               | SPC25       |
| **Iron**                      | FTL         |
|                               | ADAMTS3     |
|                               | ANGPT1      |
|                               | ANGPT1      |
|                               | C9ORF47     |
|                               | ITGA7       |
|                               | NRP1        |
|                               | NRP2        |
|                               | PROK1       |
|                               | S1PR3       |
|                               | TIMP4       |
| **Angiogenesis/invasion**     |             |
|                               | ADAMTS18    |
|                               | CLDN11      |
|                               | CREG1       |
|                               | DOK5        |
|                               | DUSP4       |
|                               | NTR2K       |
|                               | PTPR2       |
|                               | TRAF4       |
| **Proliferation**             |             |
|                               | ADAMTS18    |
|                               | CLDN11      |
|                               | CREG1       |
|                               | DOK5        |
|                               | DUSP4       |
|                               | EHF         |
|                               | GPC6        |
|                               | IFI16       |
|                               | MAP3K1      |
|                               | MAPRE2      |
| **Steroid hormone regulation**|             |
|                               | AKR1C1      |
|                               | AKR1C2      |
|                               | CPE         |
|                               | CRGM        |
|                               | CYP11A1     |
|                               | CYF2J2      |
|                               | CYF39A1     |
|                               | DIO2        |
|                               | ESR1        |
|                               | PTGER3      |
|                               | PTGFR       |
| **Tumor suppressor**          |             |
|                               | DIRAS3      |
|                               | E2F2        |
|                               | FABP3       |
|                               | LYVE1       |
|                               | SMARCB1     |
| **Carcinogenesis**            |             |
|                               | DOCK4       |
|                               | ERBB3       |
|                               | ES1         |
|                               | HLA-C       |
|                               | JAZF1       |
|                               | IGK         |
|                               | IGK         |
|                               | NBR1        |
|                               | RECK        |
| **Stress response**           | BDNF        |

development of the endometrium. MMPs have also been implicated in the adhesive, invasive, and metastatic processes involved in endometriosis [37]. Both ectopic and eutopic endometrial tissues show altered levels of MMP and TIMP expression, favoring tissue invasion and remodeling.

Tumor protein 53 (TP53) regulates the cell cycle functions as a tumor suppressor and while its role in endometriosis remains unclear, there is evidence to support its apoptotic resistance and enhanced survival of endometrial cells in endometriosis [38]. TP53 was found to be overexpressed in epithelial cells in a considerable number of endometriotic lesions [39]. However, it was found that TP53 was insignificantly upregulated in endometriosis tissue when compared with control endometrium [40].

Estrogen plays a significant role in the maintenance and chronic bleeding of endometriosis. Estrogen receptor 1 (ESR1 or ERalpha) is the dominant receptor in the adult uterus and the major mediator of estrogenic effects. It plays
a role in the hormonal deregulation and inflammation seen in this disease [41]. Steroid hormone receptors such as ESR are altered in endometriosis [42, 43].

The analysis of the combined pathway of over- and underexpressed genes, as summarized in Table 1, revealed that tissue inhibitor of metallopeptidase 1 (TIMP1) may participate in the process of invasion and tissue remodeling that is hypothesized to occur in the pathogenesis of endometriosis [44]. In endometrial carcinomas, Ephrin-B2 (EFNB2) expression may reflect or induce increased potential for growth and tumorigenicity [45].

Brain-derived neurotrophic factor (BDNF) levels are low in the follicular fluid of women with endometriosis and suggest that neurotrophins may contribute to the pathogenesis via aberrant oxidative stress mechanisms [31]. Shaco-Levy et al. (2008) found that levels of CDH1, MMP-2, and MMP-9 expressions were significantly higher in endometriosis as compared to endometrioid carcinoma, indicative of altered cell proliferation, migration, differentiation, angiogenesis, apoptosis, and host defense [46].

Increased levels of fibronectin 1 (FN1) by peritoneal macrophages in patients with endometriosis may contribute to the adhesion formation and associated reactive fibrosis seen in this disease and may influence the implantation of endometrial cells and their subsequent growth in the pelvis [47].

Phosphoinositide-3-kinase and RAS/RAF/MAPK pathways have been suggested to be involved in the initial development of endometriosis [48]. Intercellular adhesion molecule 1 (ICAM1) may play a role in the early implantation of peritoneal endometriosis [49].

4.3. Transitional Zone. While the analysis of the upper clade of the endometriotic specimens showed a particular biosignature, the analysis of the lower clade, composed of eutopic endometrial tissue of patients with endometriosis, revealed two distinct biosignatures, one specific to the lower clade and the other to the transitional zone. Although different sets of genes were identified, they are also involved in the control of inflammation, the immune response, apoptosis, cell proliferation, and lipid metabolism (Table 2). The chymase 1 gene (CMA1) found in mast cells has been shown to influence the inflammatory response by converting interleukin-1 beta into the active form, interleukin 1 [50]. The prostaglandin-endoperoxide synthase 2 gene (PTGS2/COX2) has been reported to play an important role in the inflammatory response through the production of prostaglandins [51]. Meanwhile, the cannabinoid receptor 2 (CNR2) has been shown to play an anti-inflammatory and antioxidative role in mice that have undergone chemotherapy [52].

Other studies have shown that member A of the Ras homolog gene (RHOA) can influence cell apoptosis in heart muscle cells [53]. The caspase 3 gene has also been found to induce apoptosis in cells when overexpressed [54], but possess a negative feedback mechanism as well to prevent excessive and potentially harmful mass cell death [55]. The other apoptosis-related gene is BAD; it could induce apoptosis through cleavage by caspases or inhibit apoptosis if the gene is overexpressed [56]. Finally, studies have found that the ABCA1 gene plays a major role in cholesterol transport across cell membranes [57]. This can greatly affect the synthesis of steroid hormones such as estrogen, which is well known to possess a strong stimulating effect on endometriotic growth [58]. Among the gene synapomorphies (biomarkers) identified is DRD2, which has recently been linked to eutopic and ectopic endometriotic lesions and suggested as a target to develop therapeutics [59].

4.4. Applications in Diagnosis and Prognosis. Expression profile of specimens at the border between eutopic endometrium and endometriosis specimens, interestingly, revealed that the overexpressed estrogen receptor 2 (ESR2) is a central linkage to other gene nodes. The transitional status of these specimens is highlighted by the mostly dichotomously expressed synapomorphies (Figure 4). This is an important finding because it shows that the overexpression of ESR2 could be the triggering step that initiates the deregulation of other key genes associated with inflammation, cellular matrix, immune response, growth factors, apoptosis, and others, thus leading to endometriosis (Supplementary Material 3). Indeed, several studies have reported high expression of ESR2 but lower levels of ESR1 in endometriotic tissue which caused a decrease in ESR1/ESR2 ratio [41, 60, 61] and which is in agreement with our findings (see also ESR1 in Figure 3). While Bulun and colleagues recently proposed a hypothetical model where the strikingly low ratio of ESR1/ESR2 could shift the stimulatory effect of estradiol on the progesterone receptor expression [62, 63], our study showed that the overexpression of ESR2 could precede the pathological and clinical signs of endometriosis; these potentially at-risk specimens grouped together closer to diseased specimens. The overexpression of ESR2 could be triggered by several factors ranging from genetic predisposition [64] to environmental exposures [65–67]. ESR2 polymorphism has been reported to play a role in endometriosis in various populations such as Brazilian [68] and Japanese women [69]. The disruption of ESR2 and the ensuing decrease of the ESR1/ESR2 ratio could be the culprit for the cascade of molecular events that initiates cellular deregulation and tissue remodeling associated with endometriosis (Figure 6). The screening for increased ESR2 expression could offer a diagnostic tool to identify women at risk of developing endometriosis.

It should also be noted that the endometrial tissue of women with endometriosis is different from the endometrial tissue of healthy women without the disease. For example, differences in proliferation of endometrial epithelial, stromal, and endothelial cells [70, 71], spontaneous apoptosis [72, 73], expression of cell adhesion molecules [74], and production of steroids and cytokines [74, 75] have been found. The limitation of our study is that we are restricted by the original design of published studies as deposited in the public domain. Although this dataset was limited to only 20 specimens, Hever and colleagues successfully published their findings [17] and made them available. It is important to note that endometriosis omics data available in the public domain is limited.
Table 2: Gene list summarizing the biosignature of the eutopic clade.

| Function                              | Gene symbol |
|---------------------------------------|-------------|
| Cell cycle                            | CDKN2A, PTGS2, EGFR, MDM2, F2 |
| Clotting/vascular integrity           | F2, NOS1, RHOA, MMP1, SP1 |
| Cell adhesion and migration           | CD36, F2, RHOA, JAG1, MDM2, EGFR, CDKN2A |
| Apoptosis                             | MDM2, BAT3, BCL2L1, IL17A, RARA |
| Clotting/vascular integrity           | F2, NOS1, RHOA, MMP1, SP1 |
| Clotting/vascular integrity           | F2, NOS1, RHOA, MMP1, SP1 |
| Cell adhesion and migration           | CD36, F2, RHOA, JAG1, MDM2, EGFR, CDKN2A |
| Apoptosis                             | MDM2, BAT3, BCL2L1, IL17A, RARA |
| Infertility/fertility                 | MMP1, EGFR, BRCA2, BCL2L1, PTGS2 |
| Immunity                              | CD19, CD68, IL17A, IL1RN, RARA |
| Inflammation                          | MMP1, IL17A, CMA1, CDKN2A, BMP1 |
| Cell structure                        | KCNJ5, TPS3, RHOA, EGFR, F2 |
| Carcinogenesis                        | TPM3, KLK3, EGFR |
| Angiogenesis/invasion                 | JAG1, KLK3, PTGS2 |
| Cell proliferation                    | EGFR, CDKN2A, BCL2L1, PTGS2, BCL2L1 |
| Organogenesis                         | SP1, DRD2, TPS3, JAG1, MDM2 |
| Steroid hormone regulation            | RARA, RHOA |
| Tumor suppressor                      | MDM2, BAT3, BRCA2, NOS1 |
| Stress response                       | FADS1, NOS1, RHOA |
| Protein ubiquitination                | BTRC, MDM2 |
| Lipid metabolism                     | ABCA1, FADS1, PTGS2, CD36 |
| Dopamine receptor                     | DRD2 |
| Ion transport                         | KCNJ5, SP1, RARA, DRD2, PTGS2 |
| Glucose homeostasis                   | BAD, RHOA |
| Cell debris removal                   | CD68, CD36 |
| Water and ion balance                 | DRD2 |
| Cell differentiation                  | FADS1, JAG1, SP1, RHOA |

In summary, through this study, we have shown that the biosignature of the endometriotic lesion is different from that of the endometrial eutopic tissue. Furthermore, we have revealed a particular biosignature for specimens that are in a transitional state to develop uterine endometriosis. This study contributed a novel phylogenetic approach to modeling the molecular heterogeneity of endometriosis patients into a tree-like hierarchical cladogram that reveals the simultaneously deregulated gene expressions—also termed clonal or driver aberrations. This data-based analysis shows not only directionality of change from eutopic endometrium to ectopic endometriosis, but also its usefulness in categorizing specimens according to the accumulation of molecular changes, which can be applied in diagnosing or for screening patients at-risk for developing endometriosis. In addition to supporting the ESR1 to ESR2 ratio hypothesis on the initiation of endometriosis, we have shed light on new genes and pathways that were not previously described as significant to the pathology of endometriosis. This work is a necessary first step in examining novel gene networks by a biologically compatible method that could shed light on principal drivers of the disease development process.

Glossary of Phylogenetics Terminology Used in This Paper

Since the field of phylogenetics, already extensively used in biology, zoology, botany, virology, and parasitology for over 50 years, is new to the biomedical field, we think that
Figure 6: Diagram summarizing the central role of ESR2 in triggering the molecular cascade of cell and tissue dysfunction in the transitional zone, which could lead to endometriosis.

providing a glossary would be useful to the reviewers and readers.

Clade. A group of specimens sharing one or more synapomorphies.

Cladogram. A graphic representation of relationships among specimens based on the synapomorphies (shared derived characters). The cladogram is a summary of trends that occur in the data while the upper part of it represents the specimens with highest amount of synapomorphies (shared mutations).

Dynamic Classification. A classification that has the capacity to incorporate new novel specimens without major alterations to its main groups.

Outgroup. The group of specimens used to polarize the ingroup values of gene expression into ancestral (plesiomorphic) and derived (apomorphic).

Ingroup. The group of specimens under study, for example, cancer specimens or endometriosis specimen in this study.

Parsimony. Means simplicity, the preferred hypothesis is the one requiring the least number of explanations (Occam’s Razor). In the context of our work, the preferred phylogenetic tree is the tree that requires the least number of steps to construct it from the polarized data matrix.

Polarity Assessment. Also known as outgroup comparison. It is the basis of sorting out the data values (whether proteomic (m/z), or microarray expression values) into ancestral and derived. By using our algorithms (UNIPAL/E-UNIPAL), we transform absolute numbers from data values into polarized binary numbers (0/1), where zero (0) signifies ancestral and one (1) signifies derived.

Predictive Classification. A classification that reveals the characteristics (or profile/pattern) of a specimen when its place in the classification is known.

Phylogenetic Classification [Phylogenetic Systematics]. A classification that uses synapomorphies to delimit clades (i.e., monophyletic groups).

Synapomorphy. A shared derived protein or gene expression value in comparison with a number of normal specimens (the outgroup). A protein synapomorphy may have one of the following conditions: (1) a new novel protein, (2) a disappeared protein, (3) up regulated protein, (4) down regulated protein, and (5) asynchronously regulated protein (the m/z values are above and below the normals’ range but not within the normals’ range). A gene synapomorphy may have one of the following conditions: (1) overexpressed value above normals’ range, (2) underexpressed value below the normals’ range, (3) dichotomously asynchronous values, and (4) unmeasurable expression value.

Abbreviations

UNIPAL: Universal Parsing Algorithm
DEA: Dichotomously Expressed Asynchronous genes.

Authors’ Roles

Dr. M. Abu-Asab, performed the data analysis and participated in the study design. Mr. M. Zhang drafted parts of the paper. Dr. D. Amini contributed significantly to the discussion. Dr. N. Abu-Asab offered a critical revision of the paper content. Dr. H. Amri contributed the conception, design, and interpretation of data, preparation, and submission of paper.

Acknowledgment

The authors would like to acknowledge Dr. Anthony Scialli for reviewing the paper.

References

[1] X. Gao, J. Outley, M. Botteman, J. Spalding, J. A. Simon, and C. L. Pashos, “Economic burden of endometriosis,” Fertility and Sterility, vol. 86, no. 6, pp. 1561–1572, 2006.
[2] J. E. A. Sampson, “Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity,” American Journal of Obstetrics and Gynecology, vol. 14, pp. 422–469, 1927.
[3] R. W. Steele, W. P. Dmowski, and D. J. Marmer, “Immunologic aspects of human endometriosis,” American Journal of Reproductive Immunology, vol. 6, no. 1, pp. 33–36, 1984.
Obstetrics and Gynecology International

[37] H. W. Chung, Y. Wen, S. H. Chun, C. Nezhat, B. H. Woo, and M. Polan, “Matrix metalloproteinase-9 and tissue inhibitor of metalloproteinase-3 mRNA expression in ectopic and eutopic endometrium in women with endometriosis: a rationale for endometriotic invasiveness,” *Fertility and Sterility*, vol. 75, no. 1, pp. 152–159, 2001.

[38] D. P. Braun, J. Ding, F. Shaheen, J. C. Willey, N. Rana, and W. P. Dmowski, “Quantitative expression of apoptosis-regulating genes in endometriosis from women with and without endometriosis,” *Fertility and Sterility*, vol. 87, no. 2, pp. 263–268, 2007.

[39] T. Toki and K. Nakayama, “Proliferative activity and genetic alterations in TP53 in endometriosis,” *Gynecologic and Obstetric Investigation*, vol. 50, supplement 1, pp. 33–38, 2000.

[40] P. Laudanski, J. Szamatowicz, O. Kowalczuk, M. Kuźmicki, M. Grabowicz, and L. Chyczewski, “Expression of selected tumor suppressor and oncogenes in endometrium of women with endometriosis,” *Human Reproduction*, vol. 24, no. 8, pp. 1880–1890, 2009.

[41] O. Bukulmez, D. B. Hardy, B. R. Carr, R. A. Word, and C. R. Mendelson, “Inflammatory status influences aromatase and steroid receptor expression in endometriosis,” *Endocrinology*, vol. 149, no. 3, pp. 1190–1204, 2008.

[42] K. S. Jackson, A. Brudney, J. M. Hastings, P. A. Mavrogianis, J. J. Kim, and A. T. Fazleabas, “The altered distribution of the steroid hormone receptors and the chaperone immunophilin FKBP52 in a baboon model of endometriosis is associated with progesterone resistance during the window of uterine receptivity,” *Reproductive Sciences*, vol. 14, no. 2, pp. 137–150, 2007.

[43] T. L. Rżnow, “Estrogen metabolism and action in endometriosis,” *Molecular and Cellular Endocrinology*, vol. 307, no. 1–2, pp. 8–18, 2009.

[44] K. E. Cox, M. Piva, and K. L. Sharpe-Timms, “Differential regulation of matrix metalloproteinase-3 gene expression in endometriotic lesions compared with endometrium,” *Biology of Reproduction*, vol. 65, no. 4, pp. 1297–1303, 2001.

[45] N. Takai, “Expression of receptor tyrosine kinase EphB4 and its ligand ephrin-B2 is associated with malignant potential in endometrial cancer,” *Oncology Reports*, vol. 8, no. 3, pp. 567–573, 2001.

[46] R. Shaco-Levy, S. Sharabi, D. Benharroch, B. Piura, and N. Sion-Vardy, “Matrix metalloproteinase 2 and 9, E-cadherin, and β-catenin expression in endometriosis, low-grade endometrial carcinoma and non-neoplastic eutopic endometrium,” *European Journal of Obstetrics Gynecology and Reproductive Biology*, vol. 139, no. 2, pp. 226–232, 2008.

[47] S. Kauma, M. R. Clark, C. White, and J. Halme, “Production of fibronectin by peritoneal macrophages and concentration of fibropectin in peritoneal fluid from patients with or without endometriosis,” *Obstetrics and Gynecology*, vol. 72, no. 1, pp. 13–18, 1988.

[48] S. Matsuzaki, M. Canis, C. Vuards-Barrière, O. Boespflug-Tanguy, B. Dastugue, and G. Mage, “DNA microarray analysis of gene expression in eutopic endometrium from patients with deep endometriosis using laser capture microdissection,” *Fertility and Sterility*, vol. 84, supplement 2, pp. 1180–1190, 2005.

[49] M. H. Wu, B. C. Yang, Y. C. Lee, P. L. Wu, and C. C. Hsu, “The differential expression of intercellular adhesion molecule-1 (ICAM-1) and regulation by interferon-γ during the pathogenesis of endometriosis,” *American Journal of Reproductive Immunology*, vol. 51, no. 5, pp. 373–380, 2004.

[50] H. Mizutani, N. Schechter, G. Lazarus, R. A. Black, and T. S. Kupper, “Rapid and specific conversion of precursor interleukin 1β (IL-1β) to an active IL-1 species by human mast cell chymase,” *Journal of Experimental Medicine*, vol. 174, no. 4, pp. 821–825, 1991.

[51] C. Maihöfner, S. Probst-Cousin, M. Bergmann, W. Neuhuber, B. Neundörfer, and D. Heuss, “Expression and localization of cyclooxygenase-1 and -2 in human sporadic amyotrophic lateral sclerosis,” *European Journal of Neuroscience*, vol. 18, no. 6, pp. 1527–1534, 2003.

[52] P. Mukhopadhyay, M. Rajesh, H. Pan et al., “Cannabinoïd-2 receptor limits inflammation, oxidative/nitrosative stress, and cell death in nephropathy,” *Free Radical Biology and Medicine*, vol. 48, no. 3, pp. 457–467, 2010.

[53] D. P. Del Re, S. Miyamoto, and J. H. Brown, “RhoA/Rho kinase up-regulate Bax to activate a mitochondrial death pathway and induce cardiomyocyte apoptosis,” *Journal of Biological Chemistry*, vol. 282, no. 11, pp. 8069–8078, 2007.

[54] T. Fernandes-Alnemri, G. Litwack, and E. S. Alnemri, “CPP32, a novel human apoptotic protein with homology to Caenorhabditis elegans cell death protein Ced-3 and mammalian interleukin-1 β-converting enzyme,” *Journal of Biological Chemistry*, vol. 269, no. 49, pp. 30761–30764, 1994.

[55] P. Vito, T. Ghayur, and L. D’Adamio, “Generation of anti-apoptotic presenilin-2 polypeptides by alternative transcription, proteolysis, and caspase-3 cleavage,” *Journal of Biological Chemistry*, vol. 272, no. 45, pp. 28315–28320, 1997.

[56] Y. S. So, Y. B. Chen, I. Ivamovska et al., “BAD is a pro-survival factor prior to activation of its pro-apoptotic function,” *Journal of Biological Chemistry*, vol. 279, no. 40, pp. 42240–42249, 2004.

[57] A. T. Remaley, J. A. Stonik, S. J. Demosky et al., “Apolipoprotein specificity for lipid efflux by the human ABCAI transporter,” *Biochemical and Biophysical Research Communications*, vol. 280, no. 3, pp. 818–823, 2001.

[58] S. E. Bulun, K. Zeitoun, K. Takayama et al., “Estrogen production in endometriosis and use of aromatase inhibitors to treat endometriosis,” *Endocrine-Related Cancer*, vol. 6, no. 2, pp. 293–301, 1999.

[59] E. Novella-Maestre, C. Carda, A. Ruiz-Sauri, J. A. García-Velasco, C. Simon, and A. Pellicer, “Identification and quantification of dopamine receptor 2 in human eutopic and ectopic endometrium: a novel molecular target for endometriosis therapy,” *Biology of Reproduction*, vol. 83, no. 5, pp. 866–873, 2010.

[60] J. Fujimoto, R. Hirose, H. Sakaguchi, and T. Tamaya, “Expression of oestrogen receptor-α and -β in ovarian endometrioma,” *Molecular Human Reproduction*, vol. 5, no. 8, pp. 742–747, 1999.

[61] T. Šmuc, M. R. Pucelj, J. Šinkovec, B. Hüsien, H. Thole, and T. L. Rżnow, “Expression analysis of the genes involved in estradiol and progesterone action in human ovarian endometriosis,” *Gynecological Endocrinology*, vol. 23, no. 2, pp. 105–111, 2007.

[62] S. E. Bulun, “Endometriosis,” *New England Journal of Medicine*, vol. 360, no. 3, pp. 268–279, 2009.

[63] S. E. Bulun, Y. H. Cheng, M. E. Pavone et al., “Estrogen receptor-β, estrogen receptor-α, and progesterone resistance in endometriosis,” *Seminars in Reproductive Medicine*, vol. 28, no. 1, pp. 36–45, 2010.

[64] B. Borghese, D. Iyama, D. de Ziegler, and C. Chapron, “[Endometriosis and genetics: what responsibility for the genes?],” *Journal de Gynecologie, Obstetrique et Biologie de la Reproduction* (Paris), vol. 39, no. 3, pp. 196–207, 2010.
[65] L. Cobellis, N. Colacurci, E. Trabucco, C. Carpentiero, and L. Grumetto, "Measurement of bisphenol A and bisphenol B levels in human blood sera from healthy and endometriotic women," *Biomedical Chromatography*, vol. 23, no. 11, pp. 1186–1190, 2009.

[66] M. G. Perpora, E. Medda, A. Abballe et al., “Endometriosis and organochlorinated environmental pollutants: a case-control study on Italian women of reproductive age,” *Environmental Health Perspectives*, vol. 117, no. 7, pp. 1070–1075, 2009.

[67] P. C. Huang, E. M. Tsai, W. F. Li et al., “Association between phthalate exposure and glutathione S-transferase M1 polymorphism in adenomyosis, leiomyoma and endometriosis,” *Human Reproduction*, vol. 25, no. 4, pp. 986–994, 2010.

[68] B. Bianco, D. M. Christofolini, F. A. Mafra, A. Brandes, K. Zulli, and C. P. Barbosa, “+1730 G/A polymorphism of the estrogen receptor β gene (ERβ) may be an important genetic factor predisposing to endometriosis,” *Acta Obstetricia et Gynecologica Scandinavica*, vol. 88, no. 12, pp. 1397–1401, 2009.

[69] Z. Wang, S. Yoshida, K. Negoro, S. Kennedy, D. Barlow, and T. Maruo, “Polymorphisms in the estrogen receptor β gene but not estrogen receptor α gene affect the risk of developing endometriosis in a Japanese population,” *Fertility and Sterility*, vol. 81, no. 6, pp. 1650–1656, 2004.

[70] M. Wingfield, A. Macpherson, D. L. Healy, and P. A. W. Rogers, “Cell proliferation is increased in the endometrium of women with endometriosis,” *Fertility and Sterility*, vol. 64, no. 2, pp. 340–346, 1995.

[71] A. Jurgensen, L. Mettler, N. I. Volkov, and R. Parwaresch, “Proliferative activity of the endometrium throughout the menstrual cycle in infertile women with and without endometriosis,” *Fertility and Sterility*, vol. 66, no. 3, pp. 369–375, 1996.

[72] H. M. Gebel, D. P. Braun, A. Tambur, D. Frame, N. Rana, and W. P. Dmowski, “Spontaneous apoptosis of endometrial tissue is impaired in women with endometriosis,” *Fertility and Sterility*, vol. 69, no. 6, pp. 1042–1047, 1998.

[73] R. K. Jones, R. F. Searle, and J. N. Bulmer, “Apoptosis and bcl-2 expression in normal human endometrium, endometriosis and adenomyosis,” *Human Reproduction*, vol. 13, no. 12, pp. 3496–3502, 1998.

[74] B. A. Lessey, A. J. Castelbaum, S. W. Sawin et al., “Aber rant integrin expression in the endometrium of women with endometriosis,” *Journal of Clinical Endocrinology and Metabolism*, vol. 79, no. 2, pp. 643–649, 1994.

[75] J. Sugawara, H. Yoshida, T. Fukaya, A. Yajima, and T. Murakami, “Increased secretion of hepatocyte growth factor by eutopic endometrial stromal cells in women with endometriosis,” *Fertility and Sterility*, vol. 68, no. 3, pp. 468–472, 1997.