Genetics of endometriosis: a comprehensive review

Danilo Deiana, Stefano Gessa, Michela Anardu, Angelos Daniilidis, Luigi Nappi, Maurizio N. D'Alterio, Alessandro Pontis & Stefano Angioni

To cite this article: Danilo Deiana, Stefano Gessa, Michela Anardu, Angelos Daniilidis, Luigi Nappi, Maurizio N. D'Alterio, Alessandro Pontis & Stefano Angioni (2019): Genetics of endometriosis: a comprehensive review, Gynecological Endocrinology, DOI: 10.1080/09513590.2019.1588244

To link to this article: https://doi.org/10.1080/09513590.2019.1588244

Published online: 25 Mar 2019.
Genetics of endometriosis: a comprehensive review

Danilo Deiana a, Stefano Gessa b, Michela Anardu c, Angelos Daniilidis c, Luigi Nappi d, Maurizio N. D’Alterio e, Alessandro Pontis e and Stefano Angioni a

aDepartment of Surgical Sciences, Institute of Gynaecology and Obstetrics, University of Cagliari, Cagliari, Italy; bGenetic Laboratory, Forensics Institute, Azienda Ospedaliero Universitaria, Cagliari, Italy; cDepartment of Obstetrics and Gynecology, Hippokration Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece; dDepartment of Medical and Surgical Sciences, Institute of Obstetrics and Gynaecology, University of Foggia, Foggia, Italy; eDivision of Gynecology and Obstetrics, Ospedale S. Francesco, Nuoro, Italy

ABSTRACT
This review aimed to critically evaluate the review, observational, cohort, and case–control studies performed so far in order to assess the association between endometriosis and genetics. The search strategies used included an online search of the MEDLINE database and a manual search of relevant publications and reviews. Additional reports were collected by systematically reviewing all references from the retrieved papers. Family studies have long suggested that genetic factors play a role in the etiology of endometriosis. Nevertheless, until now, studies on candidate genes have revealed inconsistent and contradictory evidence, leading to more questions rather than clear answers. It is possible that recent technological improvements in genetic evaluation could allow for a better understanding of the pathogenic mechanisms of endometriosis in the near future.

ARTICLE HISTORY
Received 16 January 2019
Accepted 25 February 2019
Published online 16 March 2019

KEYWORDS
Endometriosis; genic polymorphism; next-generation sequencing; hereditary disease

Introduction
Endometriosis is a nonmalignant disease caused by the presence of ectopic endometrial tissue [1]. It is one of the most frequent gynaecological chronic inflammatory diseases, occurring especially in reproductive age, and it often impairs quality of life and fertility [2,3]. Its heterogeneous clinical manifestations cause it to be underdiagnosed and cause studies on the disease to show contrasting data. Recent studies have proven that endometriosis has a multifactorial etiology, with possible causes being genetic, hormonal, immunological and environmental [4]. Family and twin studies have highlighted that gene susceptibility plays a central role in disease onset, development and relapse. Certain factors make genetic research in endometriosis complicated, including phenotypical heterogeneity, uncertain prevalence, invasive diagnostic methods and possible comorbidity. This review aims to provide a comprehensive and systematic tool for those approaching this problem. Review, observational, cohort, and case–control studies that have evaluated the genetics of endometriosis are herein described. Search strategies include an online search of the MEDLINE database and a manual search of relevant publications and reviews from 1980 to 2018. Keywords included: endometriosis combined with genes or genetics.

Family and twins’ studies
Simpson et al. published the first endometriosis genetics study in 1980. The study was based on 123 women with a histological diagnosis of endometriosis. The study showed that 6.9% of all first-degree relatives were affected, while the disease prevalence in the control group (i.e. first-grade female relatives of the participants’ husbands) was lower than 1% [5]. Similar results were published by Lamb and Coxhead, with a 4.9% and 9.4% disease incidence against 2% and 1.6%, respectively, in control groups [6,7].

A study of a female Norwegian population confirmed the previous data. The sample was composed of 522 women, of which 3.9% had an affected mother and 4.8% an affected sister; in the control group, only 0.6% of all first-degree relatives were affected. The study also demonstrated that severe manifestations of endometriosis were found more often among patients with a positive family history than among those without (26% versus 12%) [8].

In addition, twin studies have been used to expand our understanding of the contribution of genetics to endometriosis. Hadfield showed concordance in 14 out of the 16 monozygotic twins, though only 9 out of the 16 had the same severity grade (III–IV) [9].

Trelor et al. demonstrated a concordance ratio of 2:1 between monozygotic and dizygotic twins and a genetic risk ratio of 2.34 for endometriosis for sister. They also confirmed that 51% of the latent liability to endometriosis may be attributed to genetic factors [10]. All this evidence demonstrates that endometriosis is a hereditary disease. An increased risk in first-degree relatives (5–8%) suggests that endometriosis has a polygenic susceptibility. This high level of risk (i.e. more than 5%, such as in polygenic diseases) suggests that endometriosis is a multifactorial disease, while the presence of different polygenic patterns could explain the phenotypic heterogeneity.

Genic polymorphism and its molecular basis
Technological progress in the field of molecular biology has increased interest in genic polymorphism identification and its
involvement in endometriosis development. The choice of candidate genes is based on the hypothesis that the genes could be involved in the physiopathological molecular mechanisms of the disease (e.g., genes involved in steroidogenesis, in sex hormone receptorial activity, in inflammation and immune response, in tissue remodeling and neoangiogenesis, in metabolism regulation and in DNA reparation).

**Genes involved in steroidogenesis and sex hormone receptorial activity**

Endometriotic tissue is responsive to sexual steroid hormones, especially to estrogen and progesterone [11]. In fact, some risk factors for the development of endometriosis are early menarche, late-onset menopause and other conditions that cause extended estrogen exposure. For this reason, the dysregulation of estrogen and progesterone receptor-ligand signaling is one of the most credited hypotheses about a possible cause of endometriosis. Two estrogenic receptor isoforms (ER-α and ER-β) are codified by two different genes (ESR1 and ESR2), with different tissue-specific distributions; they can connect with different ligands (estrogen and antioestrogen) and activate different target genes’ transcription. The influence of ESR1 polymorphism was studied in different female populations (European and Asiatic) affected by endometriosis, but the studies’ results were inconsistent. In the ESR2 gene, there is a polymorphism localized in region 3'UTR, on the nucleotide 1730 (1730A>G) [12]. It has been correlated with severe endometriosis in Japanese women [13], but no association was seen in Italian and Korean populations [12,14].

Another pathogenic hypothesis is an altered tissutal sensibility to progesterone resulting from a malfunctioning of its receptors. In particular, progesterone gene polymorphism, called PROGINS (studied also for its role in ovarian carcinoma), seems to damage receptor-ligand functionality in the target tissue. In mutant subjects, receptors are less responsive to progesterone, promoting estrogenic activity and altering hormonal equilibria. PROGINS polymorphism is more frequent in women affected, but large discrepancies exist between women of different ethnic groups. PROGINS is associated with endometriosis in Italian [14], Austrian [15] and Brazilian [16] women, while it is not in Indian [17], Australian [18] and German [19] women.

**Genes involved in inflammation and immune response**

Inflammatory response processes are also involved in endometriosis physiopathology. Cytokine polymorphisms have been the subject of several studies, although differing results have been observed. In Taiwanese women, Hsieh et al. showed an association between endometriosis and some genic polymorphisms: -509C/T in the promoter region of the transforming growth factor beta 1 (TGFβ1) gene, 627A/C in the promoter region of the interleukin 10 (IL10) gene and 881T/C in the IL2 receptor-β gene [20].

A 2017 study focused on the influence of TGFβ1 on disease development and progression in relation to hypoxia. The results supported the hypothesis that TGFβ1 is involved in endometriosis through vascular endothelial growth factor (VEGF) expression regulation. In fact, TGFβ1 increased VEGF levels; it also has an additive effect upon hypoxia, increasing VEGF transcription to 87% [21].

Because of their key role in the development of an acute phase inflammatory process and in immune response regulation, tumor necrosis factor alpha (TNFα) gene polymorphisms have also been studied. Promoter-region genetic variants do not seem to be associated with an increased risk of endometriosis development in Korean, Taiwanese and Caucasian women [22].

Based on evidence from the literature, endometriosis is characterized by immune system-specific alterations. In fact, endometriotic cells have the ability to avoid immune surveillance, granting them easier implantation and growth in ectopic places. Possible reasons for this include modifications in human leukocyte antigen (HLA) gene expression, the secretion of circulating antigens competing with superficial antigens for self-nonself identification and the production of immunosuppressive factors and T-cell effectors’ proapoptotic factors [23].

Semino et al. proposed the hypothesis that ectopic endometriotic cells could express more class I major histocompatibility complex (MHC) molecules than could eutopic endometrial cells, preventing their natural killer (NK)-mediated death and facilitating their survival [24]. Endometriotic cells obtain the same results because of their ability to produce a large number of soluble adhesion molecules, such as intercellular adhesion molecule 1 (ICAM1). These stimulating molecules, involved in NK-mediated identification, compete for the same target – superficial cells’ ICAM fraction – protecting cells from the NK cell death mechanism [25]. The HLA system (which is also called MHC) is a genic complex with the largest polymorphic genic complex in the whole human genome. A 2001 study by Wang on Chinese women demonstrated the role of HLA-B genes (but not HLA-A genes) in endometriosis development [26]. In 2003, Ishi reported a significant association between HLA alleles DRB1*1403 and HLA-DQB1*0301, but not HLADPB1, on a female Japanese population [27].

**Genes involved in tissue remodeling and neoangiogenesis**

VEGF and endothelial growth factor receptor (EGFR) are molecular factors involved in neoangiogenesis regulation, tissutal remodeling and cell proliferation, all of which are essential processes in endometriotic tissue development.

VEGF induces endothelial cell proliferation, migration and differentiation. Different studies have shown the presence of high VEGF levels in peritoneal fluid and serum, and increased mRNA and protein expression in women affected by endometriosis [28]. VEGF also causes an increase in vascular permeability and protease release. In particular, metalloproteases (MMP) are enzymes that can ‘cut’ basal membrane and matrix proteins, and they are involved in cellular invasion and tissutal remodeling processes. VEGF also prevents apoptosis in some cells. In 2014, Perini et al. showed an association between endometriosis and the risk of VEGF polymorphism −1154G > A in Brazilian women [29].

EGFR is a transmembrane glycoprotein involved in growth control and in cell differentiation and motility. EGFR polymorphism +2073A/T was studied by Hsieh et al. in 122 affected women and 139 controls in Taiwan [30]. The study demonstrated an association between EGFR polymorphism +2073A/T and a high risk of disease development. However, this result was denied by a Japanese study [31].

**Genes involved in metabolism regulation**

GSTM1, GSTP1, GSTT1, NAT1, and NAT2 genes are codified for phase 2 enzymes involved in xenobiotic metabolism (as well as toxic substances such as dioxin and aromatic polycyclic hydrocarbon). A meta-analysis by Guo on the association
between gene variants of glutathione S-transferase (GST), GSTT1M1, and GSTT1 and endometriosis showed a significant increase in risk of endometriosis. Women with genotype GSTM1null and GSTT1null had an odds ratio (OR) of 1.96 and 1.77, respectively, for developing endometriosis when compared with those of other genotypes. However, some bias within the study reduces significance of this meta-analysis [32]. Many phase 1 enzymes’ genic polymorphisms (which are also involved in estrogen metabolism) have been studied (e.g. CYP1A1 6235T/C; CYP1A1 – 34T/C; and CYP19A1 microsatellite tetranucleotide [TTTA] repeat) [33,34]. However, these statistics demonstrating an association have not been confirmed by other scientific articles involving studies of other ethnic groups [16–36].

**Genes involved in DNA repairation**

Many studies have suggested that oxidative stress is involved in the physiopathology of endometriosis [37]. Reactive oxygen species’ excess DNA damage (e.g. base substitution) could explain the metaplastic nature of the disease. X-ray repair cross-complementing (XRCC)1, XRCC3 and excision repair cross-complementing (ERCC) each play a role in the DNA repairation system. Their malfunction contributes to the development of endometriotic lesions. In 2010, Attar et al. showed no significant difference in the frequencies of the genotype and polymorphic alleles of genes APE1 (Apurinic/Apyrimidinic Endonuclease 1), XRCC1, ERCC, XPG (Xeroderma Pigmentosus gene), and HOGG1 (8-Oxoguanine DNA Glycosylase 1). An increased risk seems to be associated with genotype Thr/Thr [38].

**From aimed research to ‘hypothesis-free’ studies**

All the manuscripts discussed above are based on research related to the polymorphism of genes that are likely involved in the pathogenesis of endometriosis. We suggest radically changing the angle from which this research is approached, instead using a method called the ‘hypothesis-free’ analysis of family linkage, as well as genome-wide association studies (GWASs) and the new generation technique of DNA sequencing called ‘next-generation sequencing’ (NGS). These techniques target the whole genome, without pre-selecting a particular gene or gene region and without an initial pathogenic hypothesis. This approach has the advantage of detecting unexpected molecular alteration pathways, which can be found in other diseases but often appear differently. The analysis of family linkage and the completion of GWASs represent different approaches to research but are complementary in identifying genetic variants of the alleles. The analysis of family linkage has the main goal of detecting rare variant genes in the general population that are involved in the familial aggregation of a specific disease; in contrast, the main goal of GWASs is to identify genetic variants that are common in the general population and associated with the risk of disease [39].

Studies that utilize linkage analyses, which focus on the association between a genetic marker known and gene-disease unknown, are based on an assumption of their closeness to the chromosome and the subsequent co-segregation in the family affected. In 2005, the Treloar et al. published an important linkage study of 1176 English and Australian families; in each family, there were at least two members with a surgical diagnosis of endometriosis, mainly sisters, for a total of 4985 individuals genotyped, of which 2709 were affected with endometriosis. This study showed a strong linkage with chromosome regions 10q26 and another linkage with chromosome regions 20p1330 [40].

The fine mapping of chromosome regions 10q26 – using 1984 single nucleotide polymorphisms (SNPs) from 1144 affected families and 1190 controls, by Painter et al. in 2011 – found three different meaningful associations with endometriosis [41]. The only one that was found in an independent study was rs11592737, located on gene CYP2C19 (an important candidate gene involved in the conversion of estradiol to estrone) [42]. However, this technique is less accurate if applied in a multifactorial disease such as endometriosis; it is associated with low sensibility and, given the difficulty in replicating the positive results (connected to the need for a large number of families affected, with multiple generations available for the test), the method has been made obsolete, opening the way GWASs. This variation is considered to be indicative of the region where the mutated gene that causes the disease is likely located.

The first GWAS focusing on endometriosis was conducted in 2010 using a Japanese population, with a sample of 1432 cases and 1318 controls. The study identified a significant association ($p = 5.6 \times 10^{-12}$; OR 1.44 [1.30–1.59]), with SNP rs10965235 located on gene CDKN2B-AS1 of chromosome 9 and with SNP rs16826658 ($p = 1.7 \times 10^{-6}$; OR 1.2 [1.11–1.29]) located on gene WNT4 of chromosome 1 [43]. The first gene controls certain tumor suppressors, such as CDKN2B, CDKN2A and ARF; its inactivation has been correlated with the development of endometriosis and endometrial cancer [44]. The second is an important gene involved in the development of the female reproductive system, essential for the formation of mullerian duct [45]. This has a sequence that regulates genes ESR1 and ESR2, and as such, it is now the main candidate gene potentially responsible for the development of endometriosis and ovarian cancer. This gene was also the target of a GWAS in 2016. This study, using a sample of 7090 individuals (2594 cases of endometriosis and 4496 healthy controls), found the marker with a stronger associated risk of endometriosis: SNP rs3820282, in the region of gene WNT4 [46].

In 2011, a GWAS led by the International Endogene Consortium (IEC), analyzed 3194 English and Australian women with a surgical diagnosis of endometriosis and 7060 controls [47]. The study divided the affected patients into two groups depending on the severity endometriosis (stage I–II and stage III–IV). The study found a strong and significant linkage (especially in the stage III–IV group) with two SNPs: rs1250248 ($p = 3.2 \times 10^{-8}$; OR 1.30 [1.19–1.43]), located in gene FN1 on chromosome 2 (important for cellular migration and adhesion), and rs12700667 ($p = 1.5 \times 10^{-6}$; OR 1.38 [1.24–1.53]), located in the intergenic region of chromosome 7p15 with regulatory element(s), which is upstream of three gene candidates.

However, this GWAS was rejected based on the results of studies that came after, such as that of Grechukhina in 2012, which found a high frequency of SNP rs61764370 (31% vs. 5%, case–control) in region 3’UTR of the gene Kirsten rat sarcoma viral oncogene homolog (KRAS) in women with endometriosis [48]. The gene KRAS is considered to be an important candidate gene; the activation of KRAS in lab rats causes the development of peritoneal endometriosis in 47% of cases and ovarian endometriosis–like lesions in 100% of cases. An analysis led by Luong in 2012 had denied the association risk of the gene [49]. A meta-analysis of the 2012 GWAS, conducted by IEC and a Japanese group, confirmed the association found in their previous studies and identified an association with other loci in gene GREB1 (Growth Regulating Estrogen Receptor Binding 1), involved in estrogenic regulation, VEZT (vezatin) involved in cellular growth,
migration and adhesion, ID4 (Inhibitor of DNA binding 4) found in co-suppressor genes involved in human ovarian tumors [50].

Another GWAS, conducted in 2017, was a meta-analysis and case–control study of 17,045 cases of endometriosis and 191,596 healthy controls. The study identified five new loci with significant associated risks of endometriosis ($p < 5 \times 10^{-8}$); those are genes involved via steroid hormone synthesis (FN1, CCDC170, ESR1, SYNE1, and FSHB) for a total of 16 genomic regions associated with a risk of endometriosis in one or more populations [51].

In recent years, due mainly to a reduction in the cost and time of examination, NGS has become an innovative technique for reading the DNA. Through NGS, it is possible to analyze a greater number of fragments of DNA, to obtain the sequence of many genes at the same time, straight from the whole codifying region (i.e. exome sequencing) or from the whole genome of an individual (i.e. whole genome sequencing). Rather than performing the screening of a certain number of variations, the most common ones on the whole genome, as was done in the genome arrays of the GWASs, NGS is able to detect the rarest variations in a population. In 2014, a study by Chettier et al. used the NGS technique in families affected by endometriosis; within 13 pairs of sisters affected, five shared unique genomic regions (in chromosomes 2, 5, 9, and 11), in which were found rare genetic variants of endometriosis [52]. In February 2016, Er et al. turned their attention to a molecular characterization of endometriosis related to ovarian cancer (EOC) through the targeted NGS of 409 cancer-related genes, identifying genetic variations on genes PIK3CA and ARID1A. These genes are correlated to endometriosis, with a high risk of malignant transformation [53].

The key takeaway from the varying GWASs associated with risk of endometriosis is that these genes are not codifying; however, the genes responsible for an association with endometriosis have not yet been identified. A direct approach to identifying the target genes could be to first identify varying codifying proteins that are commonly in partnership with the risk of endometriosis. International projects related to exon sequencing have identified variants that modify the corresponding protein codification (not synonymous codification, alternative splicing sites or stopping codon gain or loss), classifying them according to their frequency. The role of protein coding variants associated with endometriosis risk was directly tested by genotyping through a specific array of Illumina exome chips (Illumina) designed to allow the simultaneous capture of rare, low frequency and common frequency exonerated variants. 7164 patients with endometriosis and 210005 controls have been analyzed. The initial association between the illness and variations in codifying genes CUBN (cubilin), CIITA (Class II Major Hystocompatibility Complex Transactivator), and PARP4 (poly ADP-ribose polymerase) presented in the has not been confirmed [54]. Therefore, no tests of the existence of variations in the coding of common proteins (or whether they are low frequency and able to influence the risk of endometriosis) have yet been determined.

Using various methods to directly modify the sequence and function of protein amino-acids would have provided more specific gene targets. The probable causal SNPs were found to be situated in sequences, not codifying but involved in regulatory functions of gene expressions, to such a degree that they can influence the development and progression of the illness [55]. Considering that the GWASs have identified different susceptible loci, the greatest challenges following up on these GWASs is to understand the functional consequences of these loci. Studying the association between genetic variations and gene expression offers a way to connect the variations of risk with the corresponding gene targets. The gene expression changes in different individuals and the levels of expression of determined genes are controlled by particular regulatory variants, called ‘expression quantitative traits loci’ (eQTL), which are able to influence the abundance of specific transcriptions [56].

The various studies have identified three eQTLs able to alter the transcription of potential gene targets: LINC0039 and CDC42 in chromosome 1, CDKN2A-AS1 in chromosome 9, and VEZT in chromosome 12. Further functional studies are necessary to confirm the role of causal genes in the different loci of susceptibility [46,57,58].

Conclusions

In endometriosis pathogenesis, the contribution of genetics is well supported by many studies (Figure 1). Human genetic variability can cause a large number of mutations; these mutations are able to alter cellular and molecular mechanisms that, on different levels, are able to facilitate the development and maintenance of the illness. Genetics studies cannot provide a simple and univocal answer on the etiology of the endometriosis. Until now, studies on candidate genes have revealed inconsistent and contradictory evidence, providing more new questions than clear answers. Through GWASs, it has been possible to identify loci that can be explored and, due to NGS, rare genetic variants with a possible association to the illness have been found. Rapid technological advancements in genetics can open doors to meaningful developments in the future, in terms of understanding the molecular mechanisms of pathogenesis, the development and maintenance of endometriosis, and determining the search key for a new therapeutic target in this highly debilitating disease [59].

Disclosure statement

No potential conflict of interest was reported by the authors.


