



Article

Identification of Candidate Genes for Endometriosis in a Three-Generation Family with Multiple Affected Members Using Whole-Exome Sequencing

Carla Lintas ^{1,2,*} , Alessia Azzarà ² , Vincenzo Panasiti ^{1,3} and Fiorella Gurrieri ^{1,2}

- ¹ Research Unit of Medical Genetics, Department of Medicine, University Campus-Biomedico of Rome, 00128 Rome, Italy; v.panasiti@unicampus.it (V.P.); gurrieri@unicampus.it (F.G.)
- ² Operative Research Unit of Medical Genetics, Fondazione Policlinico Universitario Campus Bio-Medico, 00128 Rome, Italy; a.azzara@unicampus.it
- ³ Operative Research Unit of Dermatology, Fondazione Policlinico Universitario Campus Bio-Medico, 00128 Rome, Italy
- * Correspondence: c.lintas@unicampus.it

Abstract

Background: Endometriosis is a chronic inflammatory condition affecting 10–15% of women of reproductive age. Genome-wide association studies (GWASs) have accounted for only a fraction of its high heritability, indicating the need for alternative approaches to identify rare genetic variants contributing to its etiology. To this end, we performed whole-exome sequencing (WES) in a multi-affected family. **Methods:** A multigenerational family was studied, comprising three sisters, their mother, grandmother, and a daughter, all diagnosed with endometriosis. WES was conducted on the three sisters and their mother. We used the enGenome-Evai and Varelect software to perform our analysis, which mainly focused on rare, missense, frameshift, and stop variants. **Results:** Bioinformatic analysis identified 36 co-segregating rare variants. Six missense variants in genes associated with cancer growth were prioritized. The top candidates were c.3319G>A (p.Gly1107Arg) in the *LAMB4* gene and c.1414G>A (p.Gly472Arg) in the *EGFL6* gene. Variants in *NAV3*, *ADAMTS18*, *SLIT1*, and *MLH1* may also contribute to disease onset through a synergistic and additive model. **Conclusions:** We identified novel candidate genes for endometriosis in a multigenerational affected family, supporting a polygenic model of the disease. Our study is an exploratory family-based WES study, and replication and functional studies are warranted to confirm these preliminary findings.



Academic Editors: Nikolaos Machairiotis and Chi Chiu Wang

Received: 19 May 2025

Revised: 17 July 2025

Accepted: 5 August 2025

Published: 6 August 2025

Citation: Lintas, C.; Azzarà, A.; Panasiti, V.; Gurrieri, F. Identification of Candidate Genes for Endometriosis in a Three-Generation Family with Multiple Affected Members Using Whole-Exome Sequencing. *Biomedicines* **2025**, *13*, 1922. <https://doi.org/10.3390/biomedicines13081922>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: familial endometriosis; whole-exome sequencing; germinal; causative genes

1. Introduction

Endometriosis (EMS) is a chronic inflammatory condition characterized by the presence of endometrial-like tissue outside the uterus, commonly affecting the ovaries, fallopian tubes, outer uterine surface, and other pelvic organs. This ectopic tissue responds to hormonal fluctuations during the menstrual cycle, leading to inflammation, pain, and potential adhesion formation. Clinical manifestations include chronic pelvic pain, dysmenorrhea, dyspareunia, abnormal or heavy menstrual bleeding, pain during bowel movements or urination (particularly during menstruation), chronic inflammation, irregular periods, infertility, fatigue, bloating, and nausea. Symptom severity varies among individuals, with some remaining asymptomatic [1].

The exact etiology of endometriosis remains unclear, but several theories have been proposed, including retrograde menstruation, coelomic metaplasia, and immune system dysfunction. Genetic, hormonal, and environmental factors also appear to contribute to disease susceptibility. The diagnostic process is often delayed, sometimes taking years, due to the nonspecific nature of symptoms and the overlap with other gynecological or gastrointestinal disorders. Diagnosis is typically confirmed via laparoscopy, a minimally invasive surgical procedure, although non-invasive imaging techniques and biomarker research are gaining traction [1].

Endometriosis can significantly impact a patient's quality of life, both physically and emotionally. Chronic pain and infertility may lead to anxiety, depression, and social withdrawal. Thus, early recognition and a multidisciplinary treatment approach—including hormonal therapy, pain management, surgical intervention, and psychological support—are essential for optimal care and long-term outcomes [1].

Diagnosis typically involves pelvic examination, transvaginal ultrasound, or magnetic resonance imaging (MRI). Laparoscopy is also part of the diagnostic workup, even though it is an invasive procedure [2].

Treatment strategies encompass anti-inflammatory medications (e.g., ibuprofen), hormonal therapies (e.g., oral contraceptives and progestins), and surgical interventions (e.g., laparoscopic excision or hysterectomy in severe cases). These approaches aim to reduce pain, suppress ectopic endometrial tissue activity, and improve quality of life. Hormonal therapies function by altering the hormonal milieu to prevent the cyclical stimulation of endometrial-like tissue. Options include gonadotropin-releasing hormone (GnRH) agonists and antagonists, which induce a temporary hypoestrogenic state, thereby reducing lesion size and symptom severity. However, such treatments may have side effects, including bone density loss and menopausal symptoms, often necessitating “add-back” therapy [2].

Surgical treatment, especially laparoscopic excision of visible lesions, is considered in cases resistant to medical therapy or when anatomical distortion compromises fertility. Hysterectomy, with or without bilateral salpingo-oophorectomy, may be considered as a last resort for women with intractable pain who do not desire future fertility [2].

Lifestyle modifications, including stress reduction, anti-inflammatory diets rich in omega-3 fatty acids and antioxidants, and regular physical activity, may also alleviate disease severity and support overall well-being. Complementary therapies such as acupuncture, yoga, and pelvic floor physiotherapy have shown promise in symptom management. Ultimately, an individualized, multidisciplinary approach that considers patient goals, fertility desires, and comorbid conditions is essential for optimal long-term management. [2].

Despite being considered benign, endometriosis can lead to infertility, significant impairment of the quality of life due to chronic pelvic pain, and, in some cases, can progress to cancer. Epidemiological data indicate a higher prevalence of endometrial cancer among women with endometriosis compared to the general population, suggesting endometriosis as a potential risk factor [3]. Both conditions share characteristics such as cellular invasion, neovascularization, unregulated growth, estrogen dependence, apoptosis resistance, chronic inflammation, and metastatic potential [3,4]. Increased cytokine levels, including IL-1 β , IL-6, IL-28, and tumor necrosis factors like CCL2, CCL5, and VEGF, have been implicated in the progression of endometriosis to malignancy. An increased presence of macrophages and lymphocytes in peritoneal fluid during endometriosis progression has been observed, along with a documented imbalance between Th1 and Th2 cells [5]. These dysregulated processes are influenced by hormonal imbalances, notably elevated estrogen production and progesterone resistance.

Globally, approximately 10–15% of women of reproductive age are affected by endometriosis [6]. The condition is multifactorial, arising from the interplay between genetic predisposition and environmental factors [7]. While its exact etiology remains elusive, genetic, epigenetic, hormonal, oxidative stress, and immune components are critical contributors. Immune dysregulation, involving inflammatory mediators, cytokines, and immune cells, facilitates the implantation, proliferation, angiogenesis, and development of ectopic endometrial stromal cells [1,2].

Environmental risk factors include prolonged exposure to endogenous estrogens (e.g., early menarche or late menopause), hormonal fluctuations, short menstrual cycles, heavy menstrual bleeding, dietary influences, physical inactivity, gut and vaginal microbiota composition, and smoking. Notably, smoking appears to have an inverse association with endometriosis risk, though findings are inconsistent across studies [8].

The genetic basis of endometriosis is supported by various studies. Twin studies reveal a higher concordance rate among monozygotic twins compared to dizygotic twins. Familial clustering has been documented since the 1950s, with first-degree relatives of affected women exhibiting a five- to seven-fold increased risk [9–11]. Familial cases often present earlier onset and more severe symptoms than sporadic cases [12].

Genome-wide association studies (GWASs) have identified significant associations between single-nucleotide polymorphisms (SNPs) in genes involved in sex steroid pathways, including *ESR1*, *GREB1*, *FSHB*, and *CCDC170* [13]. Family-based linkage studies in Australian and British cohorts have pinpointed linkage regions on chromosomes 10q26, 7p13–15, and 20p13. Additional risk loci on 7p15.2 and 1p36.12 have been identified through GWAS. The subsequent resequencing of candidate genes has largely failed to uncover monogenic causes, with the exception of high-penetrance variants in *NPSR1* [14]. Overall, these studies have demonstrated that even in familial cases of EMS, multiple genes contribute to disease susceptibility, highlighting the polygenic nature of EMS.

Furthermore, epigenetic mechanisms play a significant role in the onset of EMS. Genes encoding enzymes involved in estrogen metabolism exhibit promoter hypermethylation, leading to reduced estrogen degradation [15,16]. Conversely, genes such as *ESR2*, which are involved in estrogen synthesis, display hypomethylated promoters, resulting in elevated estrogen levels. Additionally, the characteristic decrease in progesterone levels observed in EMS is associated with aberrant DNA methylation and histone modifications of the progesterone receptor gene (*PGR*). Recent studies have also indicated that specific microRNAs and long non-coding RNAs may be dysregulated, further contributing to increased estrogen levels in EMS [17].

Given the complex etiology of EMS and the estimated heritability at 50% based on genome-wide association studies (GWASs) and linkage analyses, we conducted whole-exome sequencing (WES) in a three-generation family affected by EMS to identify potential candidate genes associated with the disease. This approach is particularly powerful in detecting rare variants for complex disorders like endometriosis because it allows the identification of inherited, disease-associated mutations by comparing affected and unaffected family members, reducing background genetic noise. We hypothesize that rare variants co-segregating in affected family members may contribute to EMS.

2. Materials and Methods

2.1. The Multiplex Family

A multigenerational family was referred for genetic counseling due to a high incidence of endometriosis among its members (Figure 1). Three sisters (T., 51 years old; M., 48 years old; and B., 44 years old), their mother, grandmother, and the daughter of one of the sisters were all affected. Additionally, the mother was diagnosed with breast cancer at the age

of 65. On the maternal side, three uncles died of lung cancer, possibly in association with environmental risk factor exposure during their lives. The father, currently 75 years old, has been diagnosed with senile dementia. On the paternal side, an aunt had breast cancer, and her son was diagnosed with both melanoma and sarcoma at the age of 44. The paternal grandfather also had melanoma.

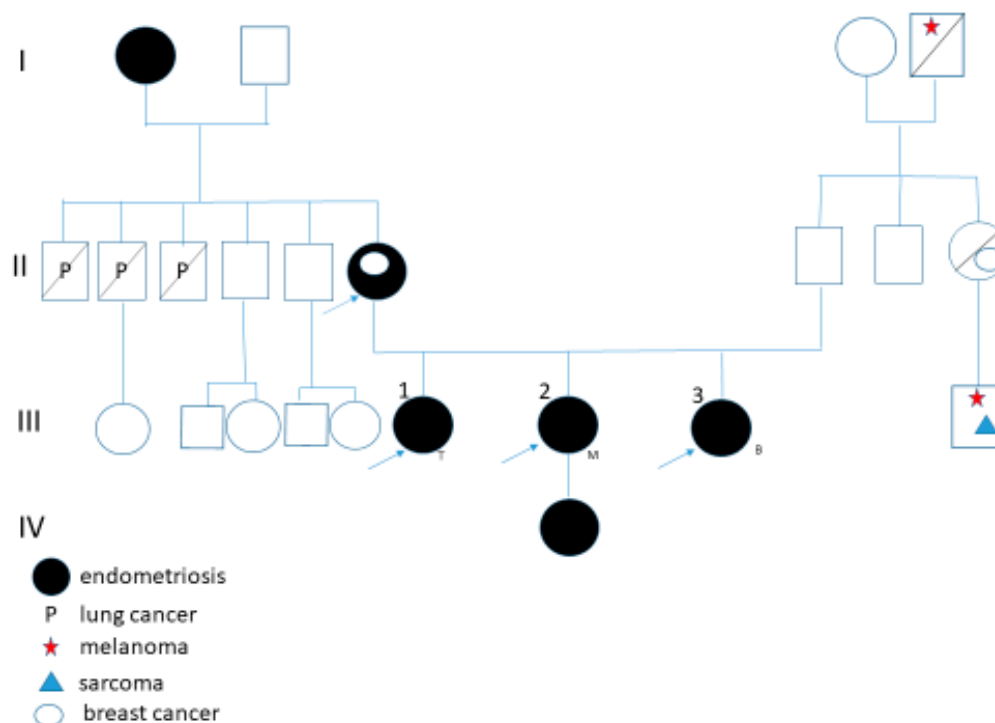


Figure 1. Pedigree of the family. The black symbols represent affected individuals by EMS. WES was performed in the women indicated by the arrow.

2.2. Whole-Exome Sequencing

Whole-exome sequencing (WES) was performed on the three affected sisters and their mother. Genomic DNA was extracted from peripheral blood leukocytes, and WES was commissioned to an external service (DanteLabs SRL, Aquila, Italy) using the Illumina platform with an average coverage of $100\times$. Details of the lab wet protocol can be found in Modi et al. 2021 [18].

2.3. Bioinformatic Analysis

FASTQ files were generated for each individual and processed using the Galaxy online platform [19] to produce individual VCF files for each proband, as well as a combined VCF file containing variants co-segregating among the four affected women. Initially, paired reads were mapped using BWA [20] (human GRCh37/hg19), followed by duplicate removal and variant calling using FreeBayes. version 1.3.7. In our WES analysis, we ensured high data quality with over 90% of the bases exceeding Q30 and coverage uniformity above 80%, supporting reliable and consistent variant detection across the exome. During our WES analysis, each individual showed $\sim 20,000$ – $25,000$ raw variants, which were reduced to $\sim 15,000$ – $20,000$ after quality filtering (e.g., depth, genotype quality, or call rate), ~ 5000 – $10,000$ after functional annotation (e.g., selecting exonic, splice-site, nonsynonymous, or LoF variants), ~ 100 – 1000 rare variants after frequency filtering (e.g., $MAF < 0.01$ or absent in population databases like gnomAD), 36 variants after applying co-segregation model, and finally 6 candidate variants after biological prioritization. Variant annotation was carried out using the enGenome-Evai software version v.0.7 [21].

Filtering criteria included coding and splicing variants with coverage ≥ 20 and a minor allele frequency (MAF) ≤ 0.01 . We selected this coverage as it ensures sufficient read depth to reliably detect heterozygous variants, minimize false positives and negatives, and improve the accuracy of variant calling, which is critical for identifying rare pathogenic mutations. MAF was set to ≤ 0.01 as we wanted to concentrate on rare variants. Initially, exome data were cross-referenced with a curated list of 49 EMS-associated genes derived from the literature [22]. Exome data were also cross-matched with GWAS data on endometriosis [13]. As no variants in the target panel were shared among the four probands, we proceeded to analyze the gene variants shared by the four probands to identify novel candidate genes.

Variant prioritization was performed using Varelect (<https://varelect.genecards.org/about/> accessed on 20 September 2024) based on a gene list derived from the analysis and incorporating the following key terms: “oxidative stress,” “inflammation,” “growth,” “adhesion,” and “estrogen.” These terms were selected from the literature. All bioinformatic analyses were conducted according to best practice guidelines [23,24]. In silico prediction tools for missense variants were accessed via Varsome, and CADD, REVEL, and PaPI scores were used to assess potential pathogenicity (<https://varsome.com/> accessed on 22 September 2024). CADD, REVEL, and PaPI are computational scores used to predict the pathogenicity of genetic variants, particularly single-nucleotide variants and missense mutations. CADD (Combined Annotation Dependent Depletion) integrates multiple annotations to estimate deleteriousness, REVEL (Rare Exome Variant Ensemble Learner) combines outputs from several tools to assess missense variant pathogenicity, and PaPI (Pathogenicity Prediction Index) uses machine learning on variant and sequence features to classify potential pathogenic effects. In variant prioritization, thresholds for the indices were set to CADD > 20 , REVEL > 0.5 – 0.7 , PaPI > 0.5 . For the in silico tools: SIFT < 0.05 (deleterious), PolyPhen-2 > 0.85 (probably damaging), and MutationTaster predictions labeled as “disease-causing” with high confidence. Variant frequencies were evaluated using the GnomAD database version 21.1 (<https://gnomad.broadinstitute.org/> accessed on 22 September 2024). GnomAD is the database related to variant frequency in the general (healthy) population.

The study was approved by the local ethics committee (Institutional Review Board approval no. 04.21) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants, and data were handled in compliance with ethical regulations.

3. Results

The flowchart (Figure 2) below summarizes the process of variant selection.

Following the completion of the bioinformatics pipeline, we excluded variants located in the 3' and 5' untranslated regions (UTRs), intronic regions, synonymous variants, and those with low quality or coverage ($\leq 20\times$). Rare copy number variants were absent.

The 14 genome-wide significant SNPs reported by Sapkota et al. [13] were not present in our exomes.

We then compared the exomes of the four affected individuals against a curated list of 49 EMS-associated genes derived from the literature [22]. No shared variants were identified within this gene panel. Subsequently, we identified 36 variants common to all four patients. These variants were prioritized using the Varelect v 5.21 software, employing the following keywords: “oxidative stress,” “inflammation,” “growth,” “adhesion,” and “estrogen.”

Each gene variant was individually evaluated based on the following:

- Minor allele frequency (MAF) and number of homozygotes, as per the GnomAD v2.1.1 database (<https://gnomad.broadinstitute.org/> accessed on 22 September 2024).
- Predicted functional impact using in silico tools such as SIFT, MutationTaster, and PolyPhen, alongside integrated scores including CADD, REVEL, and PaPI indices.
- Biological relevance to processes like growth, inflammation, adhesion, and oxidative stress, assessed through a literature review.
- Tissue-specific expression profiles, utilizing data from the GTEx database (<https://www.gtexportal.org/home/aboutAdultGtex> V10 accessed on 22 September 2024).

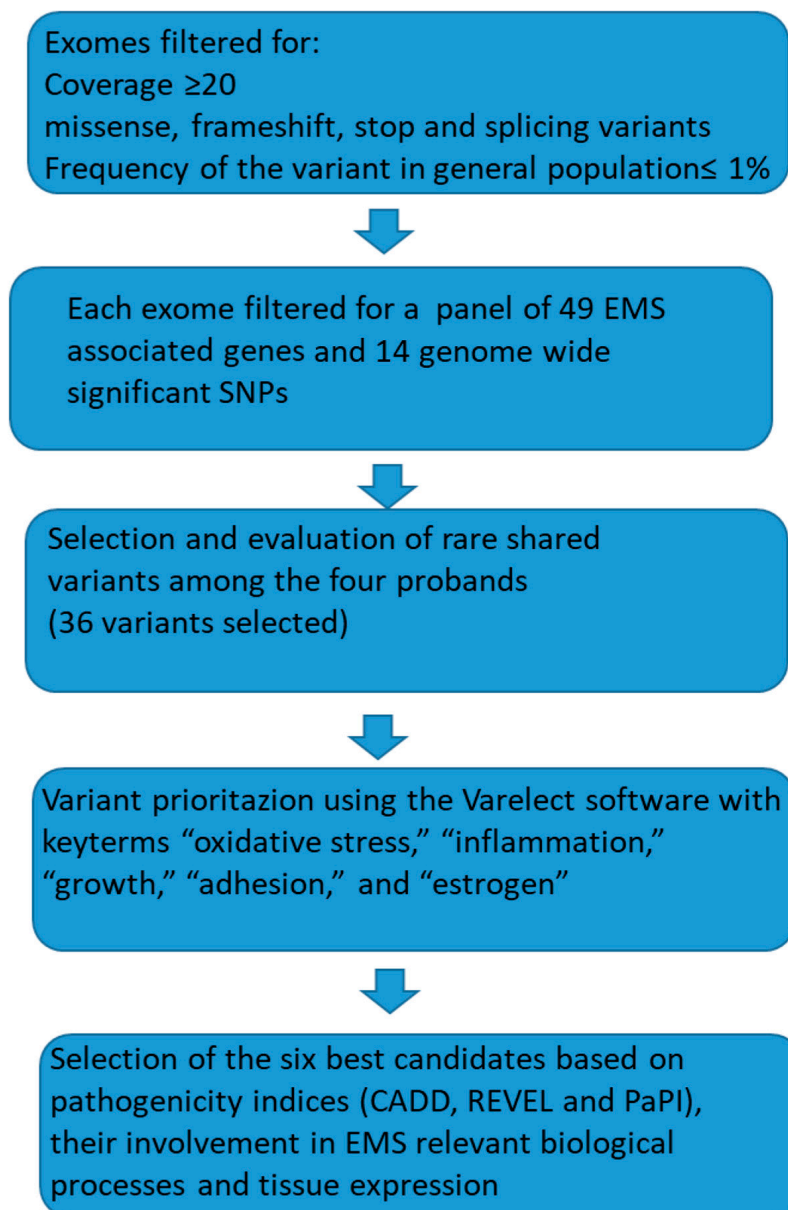


Figure 2. A summary of the workflow performed for the analysis and selection of the six best candidates. The 49 genes for the selected panel are reported in [15] and the 14 genome wide significant SNPs in [13].

Based on these criteria, we compiled a list of the most promising candidate variants (Table 1). Except for *MLH1*, none of the genes listed in Table 1 is currently associated with known clinical conditions.

The top EMS candidate gene identified is *LAMB4* (OMIM*616380), which encodes laminin beta 4. This protein is highly expressed in the skin and at lower levels in other tissues. Frameshift mutations in *LAMB4* and related laminin genes have been observed in gastric and colorectal cancers [25]. A significant proportion of these cancers with *LAMB4* mutations also exhibited loss of *LAMB4* protein expression. Laminins are crucial for maintaining normal epithelial structures and play roles in tumor invasion and metastasis, functioning as tumor suppressor genes. Germline truncating variants in *LAMB4* have been detected in individuals with colorectal cancer, with somatic second hits confirming the biallelic inactivation characteristic of tumor suppressor genes [26]. Wang et al. [27] identified another laminin gene, *LAMA4*, as an EMS-associated gene through bioinformatics analysis, while Santin et al. [22] reported *LAMA5* as an EMS candidate gene. The missense variant c.3319 G>A (p.Gly1107Arg) in *LAMB4*, detected in all four probands, has been previously reported once in the ClinVar database without association to a specific phenotype and is exceedingly rare in the general population (MAF: 3/251,316). In silico tools and high CADD, REVEL, and PaPI scores suggest this variant may impact protein function (Table 1).

The second most attractive candidate is *EGFL6* (OMIM*300239), encoding epidermal growth factor-like domain 6, expressed in cutaneous and visceral adipocytes, as well as in the vagina, cervix, breast, and lung. *EGFL6* has been shown to promote asymmetric division, maintenance, and metastasis of ALDH+ ovarian cancer cells [28]. More recently, Garrett et al. [29] demonstrated that *EGFL6* activates the MAPK pathway, enhancing cancer cell proliferation and migration in vitro, and increasing tumor growth in vivo. These experiments utilized endometrial cancer cells and tissues, with *EGFL6* knockdown resulting in reduced tumor growth. The c.1414 G>A (p.Gly472Arg) missense variant identified in the three sisters and their mother is rare in the general population (MAF: 5/183,514). In silico prediction tools and pathogenicity indices suggest this variant may function as a gain-of-function mutation, promoting abnormal growth and migration (Table 1). This variant is listed as a variant of uncertain significance (VUS) in the ClinVar database, with only one submission and no specific condition associated.

The third candidate gene, *NAV3* (OMIM*611629), functions as a tumor suppressor in breast cancer [30] and is expressed in the vagina, uterus, cervix, and other tissues. The c.7154 T>C (p.Leu2385Pro) variant is rare (MAF: 139/275,880) and has not been reported in ClinVar. In silico tools and pathogenicity indices suggest this variant could affect protein function, particularly if a second somatic mutation occurs. *NAV3* is implicated in cell migration through microtubule stabilization and inhibition of random migration [31]. Loss of *NAV3* expression has been associated with melanoma and poor survival in breast and nervous system tumors.

The *ADAMTS18* (OMIM*607512) and *SLIT1* (OMIM*603742) variants exhibit lower REVEL scores (0.088 and 0.132, respectively) and less compelling in silico predictions. However, both genes are involved in growth regulation and the identified variants are absent from the GnomAD database, suggesting to keep a potential role in EMS onset. Specifically, *ADAMTS18* has been implicated in tumorigenesis in melanoma and colorectal cancers [32], while *SLIT1* functions as a tumor suppressor in breast epithelium [33].

Table 1. List of the best candidate co-segregating variants with their associated indices and characteristics.

Gene	Gene Function	Variant	MAF	ClinVar	REVEL	CADD	PaPI	In Silico Tools: P/VUS/B+LB	CONS Score	Expression	Expression Levels in Uterine/Endometrial Tissues	Main Reference
LAMB4	Cell adhesion, growth, and migration (oncospresor)	NM_007356.3: c.3319G>A, Gly1107Arg	3/251,316	VUS (1) RCV004078131	0.715	24.6	1.0	10/9/8	4.512	Skin, but also other tissues	low	[26]
EGFL6	Growth factor (oncogene)	NM_015507.4: c.1414G>A, Gly472Arg	5/183,514	VUS (1) RCV004071274	0.367	23.5	1.0	10/9/8	5.076	Cutaneous and visceral adipocytes, vagina, cervix, breast, and lung	moderate	[29]
NAV3	Regulates microtubule dynamics, inhibits migration, and restrains dissemination of tumors (oncospresor)	NM_001024383.2: c.7154T>C, Leu2385Pro	139/275,880	/	0.425	27.8	1.0	9/7/5	7.844	Ovary, vagina, uterus, and cervix, but also other tissues	low	[30]
ADAMTS18	Proliferation and migration (oncogene)	NM_199355.4: c.1474G>T, Gly492Trp	/	/	0.088	24.1	0.92	-/5/25	3.45	Cervix, breast, cerebellum, and other tissues	moderate	[32]
SLIT1	Growth (oncospresor)	NM_003061.3: c.967T>G; Ser323Ala	/	/	0.132	23	0.98	2/6/20	5.514	Mainly brain	low	[33]
MLH1	Mismatch repair (oncospresor)	NM_000249.4: c.977T>C; Val326Ala	131/282,782	CIP (2VUS/25B-LB) RCV000035356 (most recent)	0.846	24.5	1.0	8/11/1	7.661	All tissues	high	[34]

CIP = Conflicting Interpretation of Pathogenicity for ClinVar classification, and the main accession number is reported, and in brackets number of entries. **MAF** = minor allele frequency as reported in GnomAD including allgenomic ancestry groups; **CONS** = phyloP100 conservation score; **In silico tools** = number of algorithms reprting the variant as P = pathogenic; VUS = Variant of Unknown Significance; B/LB = b/likely benign; **REVEL** (Rare Exome Variant Ensemble Learner) index is a computational score used in genetics to predict the pathogenicity; **CADD** (Combined Annotation-Dependent Depletion) score is a computational tool used in genetics to estimate the deleteriousness of genetic variants across the genome; **PaPI** (Pathogenicity Assessment of Pathogenicity Indicators) score is a computational metric used to evaluate the pathogenic potential of genetic variants, particularly rare ones; **Expression** = gene tissue expression as reported in the GTEx database (<https://www.gtexportal.org/home/> accessed 22 September 2024). In variant prioritization, commonly used thresholds include CADD > 20, REVEL > 0.5–0.7, and PaPI > 0.5. For the in silico tools: SIFT < 0.05 (deleterious), PolyPhen-2 > 0.85 (probably damaging), and MutationTaster predictions labeled as “disease-causing” with high confidence.

The last candidate gene, *MLH1* (OMIM*120436), is well known to be associated with Lynch syndrome and plays a role in the DNA mismatch repair system. The c.977 T>C (p.Val326Ala) variant, shared by all four probands, is rare (MAF: 131/282,782). High CADD, REVEL, PaPI, and pyloP100 scores, along with in silico predictions, suggest potential deleterious effects on protein function. However, multiple functional studies have demonstrated that this variant does not impair the repair system or affect splicing [34]. Consequently, the variant has been classified as benign by expert panels, with the most recent evaluation on 22 December 2024. Though the *MLH1* variant is not considered pathogenic in the context of Lynch syndrome, we cannot rule out a milder impact contributing to EMS onset, or even a second hit in the endometrial cells, due to the role of the gene *MLH1* in endometrial cancer.

4. Discussion

Endometriosis (EMS) is a multifactorial disorder arising from the interplay between environmental exposures and both common and rare genetic variants [1,6]. Among environmental factors, prolonged exposure to endogenous estrogens is paramount, followed by influences such as diet, lifestyle, and physical activity [8]. Genome-wide association studies (GWASs) and linkage analyses have identified several risk loci, including *NPSR1* (neuropeptide S receptor 1), *GREB1* (growth regulation by estrogen in breast cancer 1), *CCDC170* (coiled-coil domain-containing 170), and *ESR1* (estrogen receptor 1). Beyond genetic predisposition, emerging evidence implicates epigenetic dysregulation in EMS pathogenesis [35].

Although EMS is typically considered a benign condition, it possesses the potential for malignant transformation into endometrial, ovarian, thyroid, or breast cancers [3]. The estimated overall risk for such progression is approximately 1%, though this figure may be underestimated [36]. Notably, the candidate genes identified in our study are involved in critical cellular processes such as proliferation, adhesion, migration, and DNA repair. This observation aligns with the findings from somatic DNA analyses of endometriotic tissues, which have revealed recurrent mutations in cancer driver genes, including *ARID1A*, components of the MAPK/Ras pathway, and the PI3K-AKT-mTOR axis [37].

Our most compelling co-segregating candidate gene is *LAMB4*, which encodes laminin beta 4. Previous studies have proposed other laminin genes, such as *LAMA4* and *LAMA5*, as EMS candidates [22,26,27]. Given the tumor-suppressive functions of laminins [25], we hypothesize that the germline variant identified in our family may predispose individuals to EMS, with a subsequent somatic “second hit” potentially leading to transformation in EMS phenotype. This proposed mechanism mirrors the well-established multi-step model of familial colorectal polyposis driven by germline mutations in the tumor suppressor gene *APC*, where an initial germline mutation leads to benign polyposis, and a subsequent somatic mutation causes colorectal cancer [38].

Despite the insights provided by GWAS, linkage studies, and twin studies, a significant portion of EMS heritability remains unexplained. To address this gap, whole-exome sequencing (WES) has been employed by various research groups, including ours, to elucidate the complex genetic architecture of EMS.

Several WES studies have aimed to identify rare variants in known EMS-associated genes and to discover novel candidates. For instance, Santin et al. [22] conducted WES on 80 sporadic EMS cases, finding that 43% harbored variants in 13 candidate genes (*FCRL3*, *LAMA5*, *SYNE1*, *SYNE2*, *GREB1*, *MAP3K4*, *C3*, *MMP3*, *MMP9*, *TYK2*, *VEGFA*, *VEZT*, and *RHOJ*); 8.8% had variants in eight additional genes (*KAZN*, *IL18*, *WT1*, *CYP19A1*, *IL1A*, *IL2RB*, *LILRB2*, and *ZNF366*); and 24% carried variants in three novel genes (*ABCA13*, *NEB*, and *CSMD*). Other family-based WES studies have identified new candidate genes such as *FGFR4*, *NALCN*, and *NAV2* [39]; *TNFRSF1B*, *GEN1*, and *CRABP1* [40]; and *CIITA* [41].

Notably, variants in the first three genes were further screened in a larger EMS cohort without identifying additional carriers [39]. None of the candidate gene variants proposed in these studies co-segregated with EMS in our family. Interestingly, one of the candidate genes proposed by Nousiainen et al. [39], *NAV2*, shares significant similarity with *NAV3*, which emerged as our third most promising candidate (Table 1). The variants identified in our candidate genes are all missense mutations, a finding consistent with other studies that have predominantly reported missense variants in EMS-associated genes.

Collectively, our study and others reinforce the polygenic nature of EMS. The five co-segregating variants presented in Table 1 may contribute to EMS pathogenesis through additive or synergistic effects. Among these, *MLH1* is noteworthy; although the variant co-segregating in our family has been classified as benign by expert panels due to its lack of impact on DNA repair and splicing [34], its potential contributory role in EMS cannot be entirely dismissed, especially within a polygenic framework. Supporting this, the variant exhibits high CADD, REVEL, and PaPI scores (Table 1).

The elevated computational pathogenicity scores suggest that, while individually insufficient to cause disease, the *MLH1* variant may act as a genetic modifier in combination with other risk alleles. This aligns with the increasingly recognized model in which complex diseases such as EMS arise from the cumulative effects of multiple low-penetrance variants, each subtly impacting cellular pathways involved in inflammation, immune response, angiogenesis, or hormone regulation. In this context, the presence of high-scoring *MLH1* variants may enhance susceptibility when accompanied by other functionally relevant variants in key regulatory genes. Further functional validation, including transcriptomic and proteomic profiling in affected individuals, will be essential to delineate the biological relevance of these co-segregating variants and their role in modulating disease severity and phenotypic heterogeneity.

In summary, our study has identified five novel EMS top candidate genes co-segregating within a family comprising five affected individuals. Future objectives include (i) functional validation of the identified co-segregating variants (Table 1); (ii) conducting WES in additional families with multiple EMS-affected members to expand the list of candidate genes or confirm any of the genes in our list; (iii) conducting case-control association studies. Upon establishing an extended target panel incorporating candidate genes from other studies, we plan to screen a larger cohort to further elucidate the genetic underpinnings of EMS.

We are aware that our study is limited by the small sample size derived from a single pedigree, the absence of unaffected family controls that restricts the strength of segregation analysis, the lack of replication in unrelated endometriosis cohorts, and the absence of transcriptomic or protein-level validation to support the functional impact of the identified variants. Despite these limitations, the study represents a valid and valuable effort to identify rare, potentially pathogenic variants through a family-based WES.

Author Contributions: Conceptualization, F.G. and C.L.; methodology, C.L. and A.A.; formal analysis, C.L.; investigation, C.L.; resources, V.P.; data curation, writing and original draft preparation, C.L. and F.G.; writing—review and editing, C.L., F.G., V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was approved by the local ethics committee of Fondazione Policlinico Universitario Campus Biomedico (Institutional Review Board approval no. 04.21, 20 January 2021) and conducted in accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all the subjects involved in the study, and written informed consent has been obtained from the patients to publish this paper.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. As-Sanie, S.; Mackenzie, S.C.; Morrison, L.; Schrepf, A.; Zondervan, K.T.; Horne, A.W.; Missmer, S.A. Endometriosis: A Review. *JAMA* **2025**, *334*, 64–78. [[CrossRef](#)] [[PubMed](#)]
2. Chen, Y.; Waseem, S.; Luo, L. Advances in the diagnosis and management of endometriosis: A comprehensive review. *Pathol. Res. Pract.* **2025**, *266*, 155813. [[CrossRef](#)] [[PubMed](#)]
3. Ye, J.; Peng, H.; Huang, X.; Qi, X. The association between endometriosis and risk of endometrial cancer and breast cancer: A meta-analysis. *BMC Womens Health* **2022**, *22*, 455. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
4. Dentillo, D.B.; Meola, J.; Ferriani, R.A.; Rosa-E-Silva, J.C. Common Dysregulated Genes in Endometriosis and Malignancies. *Rev. Bras. Ginecol. Obstet.* **2016**, *38*, 253–262. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
5. Ma, R.; Zheng, Y.; Wang, J.; Xu, H.; Zhang, R.; Xie, Z.; Zhang, L.; Zhao, R. Identification of key genes associated with endometriosis and endometrial cancer by bioinformatics analysis. *Front. Oncol.* **2024**, *14*, 1387860. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
6. Giudice, L.C.; Kao, L.C. Endometriosis. *Lancet* **2004**, *364*, 1789–1799. [[CrossRef](#)] [[PubMed](#)]
7. Cousins, F.L.; McKinnon, B.D.; Mortlock, S.; Fitzgerald, H.C.; Zhang, C.; Montgomery, G.W.; Gargett, C.E. New concepts on the etiology of endometriosis. *J. Obstet. Gynaecol. Res.* **2023**, *49*, 1090–1105. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
8. Polak, G.; Banaszewska, B.; Filip, M.; Radwan, M.; Wdowiak, A. Environmental Factors and Endometriosis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 11025. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
9. Simpson, J.L.; Elias, S.; Malinak, L.R.; Buttram, V.C., Jr. Heritable aspects of endometriosis. *I. Genet. Stud. Am. J. Obstet. Gynecol.* **1980**, *137*, 327–331. [[CrossRef](#)] [[PubMed](#)]
10. Parasar, P.; Ozcan, P.; Terry, K.L. Endometriosis: Epidemiology, Diagnosis and Clinical Management. *Curr. Obstet. Gynecol. Rep.* **2017**, *6*, 34–41. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
11. Frey, G.H. The familial occurrence of endometriosis; report of five instances and review of the literature. *Am. J. Obstet. Gynecol.* **1957**, *73*, 418–421. [[CrossRef](#)] [[PubMed](#)]
12. Kennedy, S.; Hadfield, R.; Mardon, H.; Barlow, D. Age of onset of pain symptoms in non-twin sisters concordant for endometriosis. *Hum. Reprod.* **1996**, *11*, 403–405. [[CrossRef](#)] [[PubMed](#)]
13. Sapkota, Y.; Steinthorsdottir, V.; Morris, A.P.; Fassbender, A.; Rahmioglu, N.; De Vivo, I.; Buring, J.E.; Zhang, F.; Edwards, T.L.; Jones, S.; et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nat. Commun.* **2017**, *8*, 15539. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
14. Tapmeier, T.T.; Rahmioglu, N.; Lin, J.; De Leo, B.; Obendorf, M.; Raveendran, M.; Fischer, O.M.; Bafligil, C.; Guo, M.; Harris, R.A.; et al. Neuropeptide S receptor 1 is a nonhormonal treatment target in endometriosis. *Sci. Transl. Med.* **2021**, *13*, eabd6469. [[CrossRef](#)] [[PubMed](#)]
15. Chen, H.; Malentacchi, F.; Fambrini, M.; Harrath, A.H.; Huang, H.; Petraglia, F. Epigenetics of Estrogen and Progesterone Receptors in Endometriosis. *Reprod. Sci.* **2020**, *27*, 1967–1974. [[CrossRef](#)] [[PubMed](#)]
16. Rossi, M.; Seidita, I.; Vannuccini, S.; Priszano, M.; Donati, C.; Petraglia, F. Epigenetics, endometriosis and sex steroid receptors: An update on the epigenetic regulatory mechanisms of estrogen and progesterone receptors in patients with endometriosis. *Vitam. Horm.* **2023**, *122*, 171–191. [[CrossRef](#)] [[PubMed](#)]
17. Ochoa Bernal, M.A.; Fazleabas, A.T. The Known, the Unknown and the Future of the Pathophysiology of Endometriosis. *Int. J. Mol. Sci.* **2024**, *25*, 5815. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
18. Modi, A.; Vai, S.; Caramelli, D.; Lari, M. The Illumina Sequencing Protocol and the NovaSeq 6000 System. *Methods Mol. Biol.* **2021**, *2242*, 15–42. [[CrossRef](#)] [[PubMed](#)]
19. Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res.* **2022**, *50*, W345–W351, Erratum in *Nucleic Acids Res.* **2022**, *50*, 8999. <https://doi.org/10.1093/nar/gkac610>. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
20. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
21. Nicora, G.; Limongelli, I.; Gambelli, P.; Memmi, M.; Malovini, A.; Mazzanti, A.; Napolitano, C.; Priori, S.; Bellazzi, R. CardioVAI: An automatic implementation of ACMG-AMP variant interpretation guidelines in the diagnosis of cardiovascular diseases. *Hum. Mutat.* **2018**, *39*, 1835–1846. [[CrossRef](#)] [[PubMed](#)]
22. Santin, A.; Spedicati, B.; Morgan, A.; Lenarduzzi, S.; Tesolin, P.; Nardone, G.G.; Mazzà, D.; Di Lorenzo, G.; Romano, F.; Buonomo, F.; et al. Puzzling Out the Genetic Architecture of Endometriosis: Whole-Exome Sequencing and Novel Candidate Gene Identification in a Deeply Clinically Characterised Cohort. *Biomedicines* **2023**, *11*, 2122. [[CrossRef](#)]

23. Yang, H.; Wang, K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* **2015**, *10*, 1556–1566. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
24. Stelzer, G.; Plaschkes, I.; Oz-Levi, D.; Alkelai, A.; Olender, T.; Zimmerman, S.; Twik, M.; Belinky, F.; Fishilevich, S.; Nudel, R.; et al. VarElect: The phenotype-based variation prioritizer of the GeneCards Suite. *BMC Genom.* **2016**, *17* (Suppl. S2), 444. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
25. Choi, M.R.; An, C.H.; Yoo, N.J.; Lee, S.H. Laminin gene LAMB4 is somatically mutated and expressionally altered in gastric and colorectal cancers. *APMIS* **2015**, *123*, 65–71. [[CrossRef](#)] [[PubMed](#)]
26. Smith, C.G.; Naven, M.; Harris, R.; Colley, J.; West, H.; Li, N.; Liu, Y.; Adams, R.; Maughan, T.S.; Nichols, L.; et al. Exome resequencing identifies potential tumor-suppressor genes that predispose to colorectal cancer. *Hum. Mutat.* **2013**, *34*, 1026–1034. [[CrossRef](#)] [[PubMed](#)]
27. Wang, T.; Jiang, R.; Yao, Y.; Qian, L.; Zhao, Y.; Huang, X. Identification of endometriosis-associated genes and pathways based on bioinformatic analysis. *Medicine* **2021**, *100*, e26530. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
28. Bai, S.; Ingram, P.; Chen, Y.C.; Deng, N.; Pearson, A.; Niknafs, Y.S.; O’Hayer, P.; Wang, Y.; Zhang, Z.Y.; Boscolo, E.; et al. EGFL6 Regulates the Asymmetric Division, Maintenance, and Metastasis of ALDH+ Ovarian Cancer Cells. *Cancer Res.* **2016**, *76*, 6396–6409, Erratum in *Cancer Res.* **2017**, *77*, 2175. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
29. Garrett, A.A.; Bai, S.; Cascio, S.; Gupta, N.; Yang, D.; Buckanovich, R.J. EGFL6 promotes endometrial cancer cell migration and proliferation. *Gynecol. Oncol.* **2024**, *185*, 75–82. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
30. Cohen-Dvashi, H.; Ben-Chetrit, N.; Russell, R.; Carvalho, S.; Lauriola, M.; Nisani, S.; Mancini, M.; Nataraj, N.; Kedmi, M.; Roth, L.; et al. Navigator-3, a modulator of cell migration, may act as a suppressor of breast cancer progression. *EMBO Mol. Med.* **2015**, *7*, 299–314. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
31. Bugaeva, O.; Maliniemi, P.; Prestvik, W.S.; Leivo, E.; Kluger, N.; Salava, A.; Virtanen, S.; Jääntti, K.; Saksela, O.; Lehti, K.; et al. Tumour Suppressor Neuron Navigator 3 and Matrix Metalloproteinase 14 are Co-expressed in Most Melanomas but Downregulated in Thick Tumours. *Acta Derm. Venereol.* **2023**, *103*, adv00883. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
32. Wei, X.; Prickett, T.D.; Vilorio, C.G.; Molinolo, A.; Lin, J.C.; Cardenas-Navia, I.; Cruz, P.; NISC Comparative Sequencing Program; Rosenberg, S.A.; Davies, M.A.; et al. Mutational and functional analysis reveals ADAMTS18 metalloproteinase as a novel driver in melanoma. *Mol. Cancer Res.* **2010**, *8*, 1513–1525. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
33. Marlow, R.; Strickland, P.; Lee, J.S.; Wu, X.; Pebenito, M.; Binnewies, M.; Le, E.K.; Moran, A.; Macias, H.; Cardiff, R.D.; et al. SLITs suppress tumor growth in vivo by silencing Sdf1/Cxcr4 within breast epithelium. *Cancer Res.* **2008**, *68*, 7819–7827, Erratum in *Cancer Res.* **2010**, *70*, 3853. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
34. Takahashi, M.; Shimodaira, H.; Andreutti-Zaugg, C.; Iggo, R.; Kolodner, R.D.; Ishioka, C. Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. *Cancer Res.* **2007**, *67*, 4595–4604. [[CrossRef](#)] [[PubMed](#)]
35. Ducreux, B.; Patrat, C.; Firmin, J.; Ferreux, L.; Chapron, C.; Marcellin, L.; Parpex, G.; Bourdon, M.; Vaiman, D.; Santulli, P.; et al. Systematic review on the DNA methylation role in endometriosis: Current evidence and perspectives. *Clin. Epigenetics* **2025**, *17*, 32. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
36. Ioannidou, A.; Sakellariou, M.; Sarli, V.; Panagopoulos, P.; Machairiotis, N. New Evidence About Malignant Transformation of Endometriosis-A Systematic Review. *J. Clin. Med.* **2025**, *14*, 2975. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
37. Anglesio, M.S.; Yong, P.J. Endometriosis-associated Ovarian Cancers. *Clin. Obstet. Gynecol.* **2017**, *60*, 711–727. [[CrossRef](#)] [[PubMed](#)]
38. Lamlum, H.; Ilyas, M.; Rowan, A.; Clark, S.; Johnson, V.; Bell, J.; Frayling, I.; Efstathiou, J.; Pack, K.; Payne, S.; et al. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: A new facet to Knudson’s ‘two-hit’ hypothesis. *Nat. Med.* **1999**, *5*, 1071–1075. [[CrossRef](#)] [[PubMed](#)]
39. Nousiainen, S.; Kuismin, O.; Reinikka, S.; Manninen, R.; Khamaiseh, S.; Kuivalainen, M.; Terho, A.; Koivurova, S.; Niinimäki, M.; Salokas, K.; et al. Whole-exome sequencing reveals candidate high-risk susceptibility genes for endometriosis. *Hum. Genom.* **2023**, *17*, 88. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
40. Kina, B.G.; Topbas Selcuki, N.F.; Bahat, P.Y.; Usta, T.; Aydin, S.; Rahmioglu, N.; Tuncer, F.N.; Oral, E. Whole exome sequencing reveals novel candidate variants for endometriosis utilizing multiple affected members in a single family. *Mol. Genet. Genom. Med.* **2024**, *12*, e2312. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
41. Zhu, Y.; Pan, H.; Han, Y.; Li, T.; Liu, K.; Wang, B. Novel missense variant of CIITA contributing to endometriosis. *Reprod. Biomed. Online* **2022**, *45*, 544–551. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.