ISSN 2282-0930 • Epidemiology Biostatistics and Public Health - XIII^th SISMEC Congress - Vol. 20 Suppl. 1

Association Analysis of Rare Variants in Endometriosis

Erika Carbone[1], Luciano Calzari[2], Daria Marzanati[1,3], Luca Pagliardini[3], Lisett Solano Narduche[3], Ludovica Bartiromo[3], Matteo Schimberni[3], Errico Zupi[4], Massimo Candiani[3], Paola Viganò[5], Davide Gentilini[1,2]

(1) *Department of Brain and Behavioral Sciences, University of Pavia*
(2) *Bioinformatics and Statistical Genomics Unit, Istituto Auxologico Italiano IRCCS*
(3) *Obstetrics and Gynaecology Unit, IRCCS San Raffaele Scientific Institute*
(4) *Department of Molecular and Developmental Medicine, Obstetrics and Gynecological Clinic, University of Siena*
(5) *Infertility Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico*

CORRESPONDING AUTHOR: *Carbone Erika, erika.carbone01@universitadipavia.it*

INTRODUCTION

Endometriosis is a chronic, inflammatory gynecological condition characterized by the presence of endometrial-like tissue outside the uterus. It appears in various forms and can be classified based on location into several subtypes: superficial peritoneal, ovarian, deep infiltrating, extra-abdominal, and iatrogenic [1]. In recent years, advances in genetics have underscored the complex, multifactorial nature of endometriosis. Genome-wide association studies (GWAS) have identified common genetic variants linked to disease susceptibility. However, Whole Exome Sequencing (WES) has emerged as a powerful tool for uncovering rare, high-impact mutations that may explain severe or familial cases. WES not only complements GWAS findings but also opens new ways for understanding disease mechanisms, identifying biomarkers, and developing personalized therapeutic strategies.

AIM

The aim of this study is to identify and compare genetic variants present in a cohort of healthy women and women with endometriosis by using WES. The goal is to discover potential genetic biomarkers and gain a better understanding of the molecular mechanisms underlying the disease, to improve early diagnosis, prognosis, and the development of targeted treatments for women with endometriosis.

METHODS

This case-control study included 400 Italian women, 200 with laparoscopically confirmed endometriosis and 200 healthy controls. A key focus of the study was an in-depth quality control analysis: particular attention was devoted to evaluating the distribution and consistency of sequencing quality metrics—such as average coverage, read depth, genotype quality, and mapping quality—across the target regions. This thorough assessment was crucial to ensure homogeneous and comparable data across all samples. Only variants meeting stringent criteria—read depth >10, genotype quality ≥30, and mapping quality ≥40—and located in regions shared by at least 95% of the samples were retained, strictly following GATK best practices. This rigorous filtering was essential to achieve consistent variant calling across the cohort. To support this process, a comprehensive bioinformatic pipeline was developed for the processing, filtering, and statistical analysis of genetic data, ensuring high reliability and reproducibility across the entire workflow. Following quality control, we focused on rare (MAF <1%), exonic and non-synonymous variants. To assess the association between these variants and endometriosis, we applied the Sequence Kernel Association Test (SKAT) using RVTESTS [2, 3]. SKAT is a powerful, regression-based method designed to evaluate the combined effect of multiple rare variants within a gene, accommodating variants that may have effects in different directions. This approach allowed us to evaluate the cumulative effect of rare variants within each gene, increasing the power to detect associations in a genetically complex disease like endometriosis. Genes with a p-value < 0.01 from the SKAT test were considered significant and were subsequently analyzed using DAVID for functional annotation and GTEx for evaluation of their tissue-specific expression, in order to explore their potential involvement in the pathophysiology of endometriosis [4].

DOI: 10.54103/2282-0930/29351

RESULTS

After quality control, we obtained 451195 variants, of which 134113 were rare, exonic, and non-synonymous. These 134113 variants were analyzed, and the SKAT test identified 98 genes with significant association ($p < 0.01$). Functional annotation revealed enrichment in glycoprotein-related genes and those involved in immune response, cell adhesion, and metabolism. Twenty-seven candidate genes showed a higher mutation burden in cases than controls, with a case-to-control burden ratio ranging from 1.1 to 5.3. Among them, ENG, PTEN, and HLA-DPB1 were implicated in known pathogenic pathways, while novel candidates like CDHR3, CSMD3, and PLA2G3 were linked to cell adhesion and inflammation. Gene expression analysis revealed relevant tissue-specific expression in reproductive organs, supporting their potential involvement in disease pathogenesis.

CONCLUSIONS

We identified 27 genes potentially implicated in the disease's development, focusing on those involved in immune response, inflammation, and tissue remodeling. Notably, genes such as ENG, PTEN, and HLA-DPB1 play key roles in cellular processes that could contribute to endometrial tissue proliferation, immune dysregulation, and fibrosis, all of which are central to endometriosis pathogenesis. Moreover, the discovery of genes like CSMD3, CDHR3, and PLA2G3 highlighted the importance of immune regulation and cellular adhesion in the progression of endometriosis. Additionally, genes such as FMO2, FMO4, SLC2A4, and TET2 suggested that metabolic dysfunctions and epigenetic alterations may also play crucial roles in disease development. These findings support the notion that endometriosis is driven by a combination of disrupted signaling pathways, immune dysfunction, altered metabolism, and epigenetic modifications. Ultimately, further research and validation of these genetic variants could lead to the identification of novel biomarkers and treatment options for endometriosis, offering hope for improved patient outcomes.

REFERENCES

1. Horne AW, Missmer SA. Pathophysiology, diagnosis, and management of endometriosis. BMJ. 2022 Nov 14;379:e070750.
2. Wu MC, Lee S, Cai T, et al. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011 Jul 15;89(1):82-93.
3. Zhan X, Hu Y, Li B, et al. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics. 2016 May 1;32(9):1423-6.
4. Sherman BT, Hao M, Qiu J, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022 Jul 5;50(W1):W216-W221.

DOI: 10.54103/2282-0930/29351