medical sciences forum

MDPI

Proceeding Paper

Potential Candidate Gene and Underlying Molecular Mechanism Involving in Tumorigenesis of Endometriosis-associated Ovarian Cancer (EAOC) in Asian Populations †.

Rut Christine Inggriani [1*], Che-Chang Chang[2,3] and Wei-Chung Yang [2,3]

1 Graduate Institute of Biomedical Informatics, Taipei Medical University; m610110001@tmu.edu.tw
2 International Ph.D.Program for Translational Science, Taipei Medical University; ccchang168@tmu.edu.tw, vyang@tmu.edu.tw
* Correspondence: m610110001@tmu.edu.tw
3 The Ph.D.Program for Translational Science, Taipei Medical University; ccchang168@tmu.edu.tw, vyang@tmu.edu.tw
† Presented at the 3rd International Electronic Conference on Cancers, 16-30 March 2023.

Abstract: Molecular aberrations in Endometriosis were known to be associated with an increased risk of epithelial ovarian cancer (EOCs), especially endometrioid ovarian cancer (EnOC) and ovarian clear cell carcinoma (OCCC). Causal genetic evidence currently remains elusive. Integrated study of related prognostic markers will help to identify the tumorigenesis pathways in endometriosis-associated ovarian cancer (EAOC). The objective of this study was to gain a better understanding of the tumorigenesis mechanisms that occur in the endometriosis-associated genetic variation-progressed ovarian cancer risk. We found 104 overlapping genes from KEGG and GO result using WGCNA analysis. To determine whether same genes were found in one or two type of the histotypes in EAOC, we overlapped data from WES and WGCNA result and got three genes, *MYH11* (found in all histotypes), *KRT5* (found in endometriosis and OCCC), and *PDGFRA* (found in endometriosis and EnOC). Interestingly, *MYH11* and *PDGFRA* involved in the role of the actin cytoskeleton. Several proteins influence the migratory and metastatic phenotype of tumor cells, directly or indirectly, as well as myosin protein and protein Platelet-Derived Growth Factor, suggesting an explanation of tumorigenesis progression in endometriosis to ovarian cancer. This analysis has provided variants fortification for further investigation in research. With the limitation of computational study, it can still prove to be an asset for the identification and treatment of endometriosis-associated ovarian cancer diseases associated within the target gene.

Keywords: Endometriosis; ovarian clear cell carcinoma; endometrioid ovarian carcinoma; tumorigenesis; exome; gene expression

Citation: Inggriani, R.C.; Chang, C.-C.; Yang, W.-C. Potential Candidate Gene and Underlying Molecular Mechanism Involving in Tumorigenesis of Endometriosis-associated Ovarian Cancer (EAOC) in Asian Populations †. 2023, 3, x. https://doi.org/10.3390/xxxxx

Published: 16 March 2023

1. Introduction

Endometriosis, a benign condition is known for affecting the gynecology system but shows many cancer-like characteristics, including the invasion of tissues, proliferation, angiogenesis, and decreased apoptosis. There is an increased risk of ovarian cancer in woman with endometriosis, according to current epidemiological studies. Recent genetic studies suggest some links between endometriosis and ovarian cancer, indicating a possible association between the molecular diagnosis of endometriosis and the progression to ovarian cancer (endometriosis-associated ovarian cancer, EAOC) [1]. Likewise, certain genetic markers predisposing to endometriosis place an individual at an increased risk of epithelial ovarian cancer, most commonly endometrioid ovarian cancer (EnOC) and clear cell carcinoma (CCC) [1,2]. Ovarian cancer accounts for about 300,000 new diagnoses and 185,000 deaths annually, being both the eighth most common and deadly cancer in women worldwide [3]. The risk of ovarian cancer is estimated to be 1.9-fold higher when you have endometriosis, according to one systematic review and meta-analysis. The risk is even higher for clear cell ovarian cancer (3.4-fold) and endometrioid ovarian cancer (2.3-fold) [4].

Of note, endometriosis still carries a higher risk of malignancy toward ovarian cancer and their prevalence in Asian population (15% of woman) known to be higher than western population (less than 5-10%). Given this plethora of ethnic variations that exist among Asian populations [5], the epidemiological data and current trends may help updated approach to determining the risk of endometriosis to ovarian cancer in Asian populations. Therefore, it is important to focus on the EAOC phenomena in Asian. EAOC cancer genomic profiling has been performed through many sequencing technologies and yet still remains enigmatic. Some potential genetic variants might play an important role in the genetic predisposition of endometriosis lesions leading to carcinogenesis through common hallmark cancers such as mismatch repair deficiency, Wnt/B-catenin signaling pathway, PTEN/PIK3 activation, ARID1A inactivation and chromatin remodeling [6]. Therefore, it is imperative that relevant regions of the genome with variations related to disease pathology need to be analyzed.

The development of ovarian cancer has been linked to variant mutations, or alterations in the DNA sequence. According to several studies [7,8], some gene mutations are linked to endometriosis and a higher risk of ovarian cancer. In endometriosis-associated ovarian cancer, it is unclear exactly how variant mutations and mRNA expression variations are related. One theory is that the variant mutations may modify mRNA expression, which would alter gene function and raise the chance of developing ovarian cancer. Alternately, alterations in mRNA expression might be a factor in the emergence of variant mutations in ovarian cancer linked to endometriosis. The relationship between endometriosis and ovarian cancer is an important subject for research since discovering this connection can help us better understand the biology of this malignancy. More investigation is needed to completely understand the complex interactions between genetic modifications, mRNA expression variations, and the onset of endometriosis-associated ovarian cancer. In this study, we aim to uncover tumorigenesis route in endometriosis associated with ovarian cancer (EAOC) and reveal genetic variant identity based on whole exome sequencing data.

2. Materials and Method

2.1. Data Collection

To conduct this study, the literature search used the most recent 5 years of published articles. The following combinations of various search terms were used: (exome sequencing) AND (endometriosis OR endometrioma) AND (ovarian clear cell carcinoma) AND/OR (endometrioid) AND (ovary OR ovarian). Also, relevant references were evaluated from eligible studies. The report included all cohort that analyze whole exome sequencing or sequencing in exon to detect somatic mutations among EAOC groups within East Asian countries to identify ovarian cancer-associated genes and variant mutations.

Incomplete clinical pathology and mutation annotations from exome sequencing were also excluded, as were publications from sources other than East Asian nations, samples with a lower size (<10), and data that couldn't be recovered (chromosome position, cDNA change, protein change). The 137 cases in the EAOC data were divided into three groups: non-endometriosis (6 cases) endometriosis (29 cases), endometrioid ovarian cancer (22 cases) (EnOC), and ovarian clear cell carcinoma (80 cases) (OCCC). Detail selected data presented in table S1[8-12]. To reveal how SNPs encoded amino acids (nonsynonymous) can simply influence promoter activity (gene expression), mRNA stability, protein localization in subcellular, GEO (Gene expression omnibus) dataset GSE65986 and GSE7846 were used.

2.2. Data Analysis

To obtain detailed variant annotations, we imported summary datasets from whole exome sequencing results into web-based variant annotation tools VEP (variant effect predictor) [13]. VEP output result can be found in VCF or TXT format. VCF is a generic format for storing DNA polymorphism data such as SNPs, insertions, deletions, and structural variants. To predict the pathogenicity impact of each variant category, a specific set of tools must be used. VEP provides much range of algorithms to help assess the process of potential functions of a variant. Here we used PolyPhen-2, SIFT4D, Provean, CADD, GERP. GSE65986 and GSE7846 were analyzed using co-expression network in iDEP web tools 1.0 [14]. Co-expression networks are found and displayed by iDEP using the weighted correlation network analysis (WGCNA) package to compare gene expression between two groups of samples (normal and EAOC). The co-expression network is divided into modules, which are colored-coded and displayed on the gene dendogram. Each module consists of a collection of genes that are closely connected among one another. These modules are the subject of GO and KEGG enrichment investigations.

Once the harmful effect prediction variants in the exome region had been filtered, genes that overlapped between the exome and mRNA datasets were screened to further investigate the gene interest. The gene must be present in at least two histotypes (EnOC or OCCC and endometriosis) of EAOC to be identified. Afterwards, we use databases such as HPA (https://www.proteinatlas.org/) [15], cBioportal (http://cbioportal.org) [16,17], and COSMIC (https://cancer.sanger.ac.uk/) [18] to adding information about gene and variant information.

The Human Protein Atlas (HPA) project is a tool for to identifying the subcellular location and expression of a certain protein in various tissues. The comprehension of the possible effect of the exonic variations on protein function and the identification of probable disease-causing processes can both benefit from this knowledge., including the ovaries or endometrium. cBioPortal is a tool for examining cancer genomics information, including information from sizable genomic studies like The Cancer Genome Atlas (TCGA). cBioPortal can offer details on the genetic mutations and modifications that are frequently connected to endometriosis-associated ovarian cancer, as well as details on the biological networks and pathways that are affected by these changes. The COSMIC (Catalogue of Somatic Mutations in Cancer) database is a repository for details on mutations found in cancer samples. It offers details on the types, frequency, and correlation of mutations with various cancer types.

3. Result

3.1. Data Preproccessing and Normalisation

In the first step, all genes with variants in exons that meet the criteria of the study design, pertaining to whole exome sequencing or targeted sequencing in exon regions were gathered from the literature. In VEP database, we identified and annotated 17 genes in 6 non-endometrial tissues, 7341 genes in 29 ectopic endometrium tissues, 8206 genes in 16 eutopic endometrium tissues, 187 genes in endometrioid tissues, and 220 genes in clear cell tissues according to their chromosome position, alternative allele, and their HGVS information collecting from the summary of the whole exome sequencing processing data in 5 literatures.

In order to analyze mRNA expression, we gathered GEO datasets from GSE65986 and GSE7846, along with microarray raw counts. Neither dataset was different from the other, and both met the tissue criteria for normal, endometriosis, and epithelial ovarian cancer (EnOC and OCCC). We used the GPL570 platform number for human microarray count. GSE7846 consisted of 5 normal and 5 endometriosis tissue from GSE7846, while GSE65986 used 25 clear cell and 14 endometrioid tissue (Table S2).

3.2. Overlapping Exonic Variant and mRNA expression data

As a result of the literature search of whole exome sequencing, a list of total EAOC genes after applying to filter was 9, 3892, 64, and 75 for non-endometriosis, endometriosis, EnOC, and OCCC histotypes, respectively. A total of 919 genes were found in 4 modules across the entire network in iDEP 1.0 analysis using the GSE65986 and GSE7846 datasets. After enrichment, there were 499 genes recognized in the database. Those genes than were intersected with the genes from WES data. There are 104 genes that overlap between the exonic variant and mRNA level (Figure S1). The following genes have been found in at least one histotypes of EAOC. They were *MYH11* in all histotypes (endometriosis, EnOC, and OCCC), *KRT5* in only endometriosis-clear cell tissue, and *PDGFRA* in only endometriosis and endometrioid. These three genes were found significantly involved in modules blue and brown. *MYH11* and *KRT5* were found in the brown module while *PDGFRA* was found in the blue module only (Figure S1b).

Using the module in iDEP 1.0, we found that *PDGFRA* is significantly related to biological processes and cellular components (ECM). *KRT5* and *MYH11* were also involved in the development of skeletal and tissue components of biological processes and found in part of the ECM (Figure S2). However, to further identify their pathway in disease, the 104 genes that overlapped were re-analyzed in DAVID using KEGG and GO enrichment. *PDGFRA* is known to be significantly related PI3K-Akt signaling pathway, focal adhesion, regulation of actin cytoskeleton, EGFR tyrosine kinase inhibitor resistance, and pathways in cancer according to KEGG. Furthermore, *KRT5* was not found in the KEGG pathway and *MYH11* involved in the Regulation of the actin cytoskeleton same as *PDGFRA* (Figure S3).

3.3. Gene of Interest for Further Investigation

MYH11, *KRT5*, *PDGFRA* were then investigated using HPA, cBioportal, and COSMIC databases. After knowing the list of genes of interest, the database was used for investigating the potential impact of protein disease-associated mechanisms. First was KRT5 (Keratin 5) which was found highly expressed protein in the female reproduction system, especially in the vagina, cervix, and breast tissue. Even though their protein was not found in the endometrium or ovary tissue, their RNA expression level was 2.0 and 2.2 normalized TPM in both tissues respectively according to HPA database. They also found highly expressed in smooth muscle cell endometrium sample among females above 30 years old. In the ovary tissue, they expressed in stromal cells for females around 50 years old and later. *MYH11* (Myosin heavy chain 11) protein was found to be highly expressed in the endometrium, breast, and ovary tissue. Their protein and RNA were equally highly expressed in the smooth muscle tissue part of the endometrium and ovary. In *PDGFRA* (Platelet-derived growth factor receptor alpha) gene, their protein, and RNA were expressed in contrast conditions. Protein in ovary tissue was lower but higher for their RNA level, in contrast with their RNA level that showed higher value in protein but lower in their endometrium tissue.

Using cbioportal database, the genetic alteration caused by *PDGFRA* (1.6%) and MYH11 (0.7%) was found significant based on GENIE cohort studies (646 samples) in OCCC and EnOC patients. Only KRT5 did not show any mutation frequency in the database. Therefore to provide complete information about somatic mutation in both genes, COSMIC database was used. For the MYH11 gene, the same variant (c.323G>A) was found in both eutopic endometriosis and endometrioid samples as well as the COSMIC database. All PDGFRA variants in our data were also found in the COSMIC database (c.2942G>A, c.2671C>T, c.2450G>A, c.2566T>G). Only KRT5 variants have not been found in the database.

4. Discussion

Extracellular matrix (ECM) plays a significant role in the development and progression of endometriosis-associated ovarian cancer (EAOC). It is possible that changes in the structure and composition of the ECM could contribute to the development of these disorders, which is a network of proteins and polysaccharides that provides structural support and signaling signals to cells. Endometriosis develops and maintains endometrial implants outside the uterus by regulating the ECM. The ECM participates in the construction of cellular structures that support the growth of malignancy in ovarian cancer, as well as the control of cell survival and proliferation. Changes in the composition and organization of ECM elements such laminin and tenascin, which have been specifically seen in ovarian cancer, are believed to have a role in the creation of the tumor microenvironment and the progression of malignancy [19].

In both endometriosis and ovarian cancer, the ECM acts as a physical and biochemical barrier to drug delivery, and altering it may enhance delivery and efficacy. Further, since alterations in the ECM have been shown to contribute to both diseases' onset and progression, targeting the ECM has been proposed as a treatment strategy. Endometriosis-associated ovarian cancer (EAOC) may be associated with alterations in genes including MYH11 and PDGFRA in the extracellular matrix (ECM), which plays an integral role in the tumor microenvironment.

The heavy chain of myosin IIA, a crucial part of the ECM and implicated in its organization and mechanical characteristics, is encoded by the *MYH11* gene. The expression of *MYH11* is often upregulated in cancer cells, which promotes cell migration and invasion by forming actin-myosin contractile structures. Consequently, cancer cells are able to move through the ECM and invade surrounding tissues, resulting in the spread of the disease. As a consequence of changing the mechanical properties of the ECM, changing cell behavior, and fostering malignancy, MYH11 mutations may play a role in ECM-mediated carcinogenesis in the setting of EAOC [20].

The protein Platelet-Derived Growth Factor Receptor Alpha (PDGFRA) plays a vital role in controlling the growth and survival of cells in endometriosis and ovarian cancer. In endometriosis, a higher expression of PDGFRA is associated with the expansion and growth of endometrial tissue outside of the uterus, resulting in the creation of endometriotic growths. This heightened expression of PDGFRA triggers the release of various signaling substances that stimulate cell growth and survival, contributing to the progression of endometriosis. Similarly, PDGFRA is often over-expressed in ovarian cancer, and its activation has been shown to drive the growth and survival of cancer cells. PDGFRA signaling pathways trigger the production of various growth factors and cytokines that contribute to the advancement of ovarian cancer [21].

Based on our finding, most of the mutation were G>A type that is known as silent mutation. Silent mutations do not alter the protein's sequence, but they can still affect how genes are expressed and cause disease even when they do not alter the protein's sequence. For instance, silent mutations can modify the secondary structure of the mRNA, which will affect how well it can be translated into protein, or they can affect the stability of the mRNA, which will vary the amount of protein generated. In some circumstances, silent mutations can also change how genes are spliced, leading to the generation of various proteins with various roles. The extracellular matrix (ECM) may change as a result of these alternative splice variants, which may also contribute to carcinogenesis in endometriosis-associated ovarian cancer (EAOC) (Figure S4).

5. Conclusion

In spite of the fact, that the function variant mutations in EAOC are still poorly understood, recent studies have indicated that they can contribute to a number of illnesses, including ovarian cancer. The effects of those mutations, including specific G>A alterations in MYH11 and PDGFRA, on the onset of EAOC require further research. Genetic variants need to be genotyped as part of further investigation. Even with the limitations of computational research, it can still be useful for locating and treating disorders linked to the target gene that are related with endometriosis-associated ovarian cancer.

References

1. Mortlock, S.; Corona, R.I.; Kho, P.F.; Pharoah, P.; Seo, J.H.; Freedman, M.L.; Gayther, S.A.; Siedhoff, M.T.; Rogers, P.A.W.; Leuchter, R.; et al. A multi-level investigation of the genetic relationship between endometriosis and ovarian cancer histotypes. *Cell Rep Med* **2022**, *3*, 100542, doi:10.1016/j.xcrm.2022.100542.
2. Murakami, R.; Matsumura, N.; Brown, J.B.; Higasa, K.; Tsutsumi, T.; Kamada, M.; Abou-Taleb, H.; Hosoe, Y.; Kitamura, S.; Yamaguchi, K.; et al. Exome Sequencing Landscape Analysis in Ovarian Clear Cell Carcinoma Shed Light on Key Chromosomal Regions and Mutation Gene Networks. *Am J Pathol* **2017**, *187*, 2246-2258, doi:10.1016/j.ajpath.2017.06.012.
3. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2018**, *68*, 394-424, doi:10.3322/caac.21492.
4. Nagle, C.M.; Olsen, C.M.; Webb, P.M.; Jordan, S.J.; Whiteman, D.C.; Green, A.C. Endometrioid and clear cell ovarian cancers: a comparative analysis of risk factors. *Eur J Cancer* **2008**, *44*, 2477-2484, doi:10.1016/j.ejca.2008.07.009.
5. Yen, C.F.; Kim, M.R.; Lee, C.L. Epidemiologic Factors Associated with Endometriosis in East Asia. *Gynecol Minim Invasive Ther* **2019**, *8*, 4-11, doi:10.4103/gmit.Gmit_83_18.
6. Shih, I.-M.; Wang, Y.; Wang, T.-L. The Origin of Ovarian Cancer Species and Precancerous Landscape. *The American Journal of Pathology* **2021**, *191*, 26-39, doi:https://doi.org/10.1016/j.ajpath.2020.09.006.
7. Anglesio, M.S.; Papadopoulos, N.; Ayhan, A.; Nazeran, T.M.; Noë, M.; Horlings, H.M.; Lum, A.; Jones, S.; Senz, J.; Seckin, T.; et al. Cancer-Associated Mutations in Endometriosis without Cancer. *New England Journal of Medicine* **2017**, *376*, 1835-1848, doi:10.1056/NEJMoa1614814.
8. Suda, K.; Nakaoka, H.; Yoshihara, K.; Ishiguro, T.; Tamura, R.; Mori, Y.; Yamawaki, K.; Adachi, S.; Takahashi, T.; Kase, H.; et al. Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. *Cell Reports* **2018**, *24*, 1777-1789, doi:10.1016/j.celrep.2018.07.037.
9. Li, X.; Zhang, Y.; Zhao, L.; Wang, L.; Wu, Z.; Mei, Q.; Nie, J.; Li, X.; Li, Y.; Fu, X.; et al. Whole-exome sequencing of endometriosis identifies frequent alterations in genes involved in cell adhesion and chromatin-remodeling complexes. *Human Molecular Genetics* **2014**, *23*, 6008-6021, doi:10.1093/hmg/ddu330.
10. Lapke, N.; Chen, C.-H.; Chang, T.-C.; Chao, A.; Lu, Y.-J.; Lai, C.-H.; Tan, K.T.; Chen, H.-C.; Lu, H.-Y.; Chen, S.-J. Genetic alterations and their therapeutic implications in epithelial ovarian cancer. *BMC Cancer* **2021**, *21*, 499, doi:10.1186/s12885-021-08233-5.
11. Yang, Q.; Zhang, C.; Ren, Y.; Yi, H.; Luo, T.; Xing, F.; Bai, X.; Cui, L.; Zhu, L.; Ouyang, J.; et al. Genomic characterization of Chinese ovarian clear cell carcinoma identifies driver genes by whole exome sequencing. *Neoplasia* **2020**, *22*, 399-430, doi:10.1016/j.neo.2020.06.002.
12. Kim, S.I.; Lee, J.W.; Lee, M.; Kim, H.S.; Chung, H.H.; Kim, J.W.; Park, N.H.; Song, Y.S.; Seo, J.S. Genomic landscape of ovarian clear cell carcinoma via whole exome sequencing. *Gynecol Oncol* **2018**, *148*, 375-382, doi:10.1016/j.ygyno.2017.12.005.
13. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biology* **2016**, *17*, 122, doi:10.1186/s13059-016-0974-4.
14. Ge, S.X.; Son, E.W.; Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **2018**, *19*, 534, doi:10.1186/s12859-018-2486-6.
15. Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Proteomics. Tissue-based map of the human proteome. *Science* **2015**, *347*, 1260419, doi:10.1126/science.1260419.
16. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **2013**, *6*, pl1, doi:10.1126/scisignal.2004088.
17. AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov* **2017**, *7*, 818-831, doi:10.1158/2159-8290.Cd-17-0151.
18. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* **2018**, *47*, D941-D947, doi:10.1093/nar/gky1015.
19. Cho, A.; Howell, V.M.; Colvin, E.K. The Extracellular Matrix in Epithelial Ovarian Cancer - A Piece of a Puzzle. *Front Oncol* **2015**, *5*, 245, doi:10.3389/fonc.2015.00245.
20. Li, Y.-R.; Yang, W.-X.; Yang, W.-X. Myosins as fundamental components during tumorigenesis: diverse and indispensable. *Oncotarget* **2016**, *7*, 46785-46812, doi:10.18632/oncotarget.8800.

21. Wilczyński, J.R.; Szubert, M.; Paradowska, E.; Wilczyński, M. Endometriosis Stem Cells as a Possible Main Target for Carcinogenesis of Endometriosis-Associated Ovarian Cancer (EAOC). *Cancers* **2023**, *15*, 111.