VGLL3 expression is associated with macrophage infiltration and predicts poor prognosis in epithelial ovarian cancer

Razaul Haque
Samsung Medical Center

Jaebon Lee
Sungkyunkwan University

Joon-Yong Chung
National Cancer Institute, National Institutes of Health

Ha-Yeon Shin
Yonsei University College of Medicine

Hyosun Kim
Yonsei University College of Medicine

Jae-Hoon Kim
Yonsei University College of Medicine

Jae Won Yun
Veterans Medical Research Institute, Veterans Health Service Medical Center

Eun-Suk Kang (eskang@skku.edu)
Samsung Medical Center

Research Article

Keywords: High-grade serous ovarian carcinoma, Hippo pathway, VGLL3, prognosis, macrophage

Posted Date: December 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2337749/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background

High-grade serous ovarian carcinoma (HGSOC) is the most common histologic type of epithelial ovarian cancer (EOC). Due to its poor survival outcomes, it is essential to identify novel biomarkers and therapeutic targets. The hippo pathway is crucial in various cancers, including gynaecological cancers. Herein, we examined the clinicopathological significance of the key genes of the hippo pathway in HGSOC.

Method

The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) data were curated to analyse the mRNA expression as well as the clinic-pathological association and correlation with immune cell infiltration in HGSOC. The protein levels of significant genes in the HGSOC tissue were analysed using Tissue Microarray (TMA)-based immunohistochemistry. An overall survival analysis was conducted using the log-rank test method and cox regression analysis. Finally, DEGs pathway analysis was performed to identify the signalling pathways associated with VGLL3.

Result

VGLL3 mRNA expression was significantly correlated with both advanced tumour stage and poor overall survival (OS) ($p = 0.046$ and $p = 0.003$, respectively). The result of IHC analysis also supported the association of high VGLL3 protein with poor OS in HGSOC. Further, VGLL3 expression was significantly associated with tumour infiltrating macrophages. VGLL3 expression and macrophages infiltration were both found to be independent prognostic factors ($p = 0.003$ and $p = 0.024$, respectively) for HGSOC. VGLL3 was associated with 3,981 gene expressions ($p < 2.04e^{-4}$), and with four known and three novel cancer-related signalling pathways, thus implying that VGLL3 is involved in the deregulation of many genes and pathways.

Conclusion

Our study revealed that VGLL3 may play a distinct role in clinical outcomes and immune cell infiltration in patients with HGSOC and that it may be a potential prognostic marker of EOC.

Introduction

Epithelial ovarian cancer (EOC) is the most dominant form of ovarian cancer, which accounts for the highest rates of mortality and morbidity among the female sex [1]. Despite the immense advancements that have been achieved in treatment strategy, EOC remains the most fatal gynaecological cancer, with a
5-year overall survival rate that has been reported to be under 45% [2]. Moreover, around 85% of EOC patients develop recurrence after first line therapy, which becomes more incurable and for which the survival rate drops down to less than 24 months [3]. Such poor prognosis may result from the complex and obscure pathogenesis of EOC, late diagnosis, a lack of predictive biomarkers, and ineffective target identification for therapeutic purposes [4]. Over the past few years, studies have elucidated different morphopathological and molecular characteristics of EOC. At present, EOC is classified into at least five categories, including high-grade serous ovarian carcinoma (HGSOC), clear cell carcinomas, endometrioid carcinomas, low-grade serous ovarian carcinomas (LGSOC) and mucinous carcinomas [5, 6]. HGSOC is the most common histological type of EOC diagnosed at advanced stages, and it contributes to an even higher percentage of mortality in ovarian cancer [7]. However, it is very difficult to characterise HGSOC, as they account for a very low number of mutations, and there is a scarcity of appropriate diagnosis markers. According to The Cancer Genome Atlas (TCGA), the only commonly mutated gene found in HGSOC is TP53 (96% cases), whereas mutations in other commonly mutated oncogenes such as KRAS, BRAF, NRAS, and PIK3CA are very rare in HGSOC (all less than 1%) [8, 9]. Meanwhile, other histological types of EOC have more mutations that frequently occur on those genes. Therefore, there is an urgent need to explore and identify novel biomarkers to improve the diagnosis and early detection of HGSOC to improve treatment efficiency.

The hippo pathway is a critical regulator of morphogenesis, organ size determination, and tumourigenesis in many tissues, including the reproductive system [10–12]. Formation of the YAP/TAZ-TEADs complex serves as the key mechanism that stimulates the expression of target genes (e.g., connective tissue growth factor and cysteine-rich angiogenic inducer 61) that are essential for cell proliferation and survival [13]. As a tumour suppressor pathway, dysregulation of the hippo pathway has been linked to various cancers [14–16]. YAP/TAZ plays an oncogenic role in EOC tumourigenesis by increasing cell proliferation and apoptotic resistance, reducing contact inhibition, and improving motility and anchorage-independent growth [17, 18]. However, the relationship between YAP/TAZ expression and clinicopathological outcomes in EOC is still a matter of debate.

Vestigial-like 3 (VGLL3) is a member of the VGLL family that also serves as a cofactor for transcriptional enhanced associate domains (TEADs). VGLL3 was first identified in humans based on its homology with the Drosophila gene vg (vestigial), where it encodes a cofactor of Scalloped, a homolog of the transcription factor TEF-1 [19, 20]. Therefore, it has been suggested that VGLL3 may function in a manner similar to vg. VGLL3 has recently been reported to be associated with the inhibition of adipocyte differentiation and the regulation of trigeminal nerve formation and cranial neural crest migration [21, 22]. Although the physiological role of VGLL3 is unknown, there is emerging evidence suggesting that VGLL3 is associated with different cancers including soft tissue carcinoma, breast cancer, lung cancer, gastric cancer, stomach adeno carcinoma, and other malignancies [23–26]. VGLL3-TEAD complex has been found to enhance cancer cell proliferation and cancer development [23–26]. VGLL3 has previously been reported to play a relevant role in EOC by Gambaro et al. in 2013 [27]. Since then, there have been no further reports elucidating the role of VGLL3 in EOC. Moreover, Gambaro et al. derived the hypothesis from a chromosome transfer experiment wherein the transfer of a chromosome fragment containing
VGLL3 gene suppressed tumour phenotypes in the ovarian tumour cell line OV90 [28, 29]. However, VGLL3 as a single-gene transfer failed to both generate stable VGLL3 expression and suppress the proliferation of OV90 cells [27]. Therefore, it is essential to re-evaluate the role of VGLL3 in EOC. Moreover, the role of VGLL3 in the immune microenvironment of EOC remains to be elucidated.

There is increasing evidence suggesting that immune cell infiltration both plays a crucial role in the prognosis of various tumours and affects OS [30–32]. Infiltration of different immune cells, such as T cells, macrophages, mast cells, and natural killer cells, is known to be associated with either favourable or unfavourable prognosis [33]. Zhang et al. recently showed that VGLL3 serves as a novel unfavourable prognostic biomarker in stomach adenocarcinoma and correlates with immune evasion, particularly due to infiltration of macrophages and dendritic cells through the analysis of mRNA expression in public databases [26].

In this study, we performed a comprehensive analysis using public databases and web tools, as well as tissue microarray (TMA) to investigate the significance of key genes in hippo pathways, including VGLL3, VGLL4, TEAD3, TEAD4, YAP, and TAZ, on the clinico-pathological characteristics and immune cell infiltration features of HGSOC. We also investigated the role of VGLL3 as a prognostic factor of HGSOC and its association with cancer-related signal transduction pathways.

**Materials And Method**

**mRNA data sources and clinical information**

The information on the mRNA expression and the clinical data of ovarian cancer were acquired from the Cancer Genome Atlas (TCGA) repository, Genome Data Analysis Centre (GDAC) (https://gdac.broadinstitute.org/), and the University of California Santa Cruz browser (https://xenabrowser.net/datapages/). For further analysis, 303 samples that had clinical and mRNA information available were selected (Table 1). The expression of genes was compared between normal ovarian (n = 88) and tumour (n = 426) tissues based on the GEPIA2 database (http://gepia2.cancer-pku.cn/#index) [34]. For the validation of gene expression in HGSOC, we acquired data from the gene expression omnibus (GEO) in the form of GSE26712 (10 normal ovarian vs. 185 tumour tissue) and GSE9891 (264 HGSOC tissue samples). RPKM and normalised read count units were used for data obtained through RNA-seq, and log2 RMA signal units were used for data obtained through expression microarray.
Table 1
Clinicopathological features of high-grade serous ovarian carcinoma patients from The Cancer Genome Atlas datasets

<table>
<thead>
<tr>
<th>Clinical Factor</th>
<th>TCGA (n = 302)</th>
<th>TMA (n = 84)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years, mean ± SD)</td>
<td>59.10 ± 10.93</td>
<td>54.77 ± 10.41</td>
</tr>
<tr>
<td>CTx response (n) (Sensitive/Resistant/unknown)</td>
<td>200/40/62</td>
<td>67/15/2</td>
</tr>
<tr>
<td>Death/Alive (n)</td>
<td>183/119</td>
<td>50/34</td>
</tr>
<tr>
<td>Overall survival (years, mean ± SD)</td>
<td>3.04 ± 2.44</td>
<td>6.65 ± 5.03</td>
</tr>
<tr>
<td>Pre-CA125 (n) (Negative ≤ 35U/ml)/(Positive &gt; 35U/ml)</td>
<td>NA</td>
<td>7/77</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage I (n)</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Stage II (n)</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>Stage III (n)</td>
<td>240</td>
<td>58</td>
</tr>
<tr>
<td>Stage IV (n)</td>
<td>38</td>
<td>13</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1 (n)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>G2 (n)</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>G3 (n)</td>
<td>260</td>
<td>44</td>
</tr>
<tr>
<td>Others (GB, GX) (n)</td>
<td>8</td>
<td>-</td>
</tr>
</tbody>
</table>

(n): Number of patients, CR: Complete response, PR: Partial response, SD: Stable disease, PD: Progressive disease, CTx: Chemotherapy, Others (GB, GX), GB: Grade borderline, GX: Grade cannot be assessed, NA: not available

Cell Culture

Nine human ovarian EOC cell lines were used: SKOV3, OVCAR3, OVCA429, OVCA433, YDOV-13, YDOV-139, YDOV-157, YDOV-161, and YDOV-151. YDOV-13 (originated from a malignant Brenner tumour), YDOV-139, YDOV-157, YDOV-161 (originated from serous cystadenocarcinomas) and YDOV-151 (originated from a mucinous cystadenocarcinoma) were the established primary cells that were used in this study [35–38]. All cell lines were kindly provided by Jae-Hoon Kim (Gangnam Severance Hospital, Yonsei University). SKOV3 and OVCAR3 cell lines were maintained in RPMI-1640 media supplemented with 10% FBS and 1% with penicillin / streptomycin. The other cell line was maintained in DMEM media containing 10% FBS and 1% penicillin / streptomycin. All the cell lines were cultured at 37°C in 5% CO₂.
Rna Isolation And Real-time Pcr

At 70–80% of confluence, all cells were washed two times with PBS, after which total RNA was extracted using TRizol reagent (Ambion, Carlsbad, USA) according to the manufacturer’s protocol. Total RNA (1 µg) from each sample was reverse-transcribed into cDNA using the Maxima First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA) according to the manufacturer’s protocol. Real-time polymerase chain reaction (PCR) was performed to quantify mRNA expression using SYBR Green PCR Master Mix (Enzymomics, Daejeon, Republic of Korea) and the QuantStudio 6 Flex real-time PCR system (Applied Biosystems, Foster City, CA) while following the manufacturer’s instructions. Relative mRNA expression was quantified using the comparative Ct (ΔCt) method and expressed as 2^−ΔΔCt. Each assay was done in triplicate and expressed as mean ± standard error (SE). To determine the reaction efficiencies, a series of dilutions was prepared from a stock solution of total RNA to generate a standard curve. The following primers were used for PCR: VGLL3: Forward 5'- CCAACTACAGTCACCTCTGCTAC-3' and Reverse 5'- ACCACGTTGATTCTTACTCTTG-3', GPADH: Forward 5'- ATGGAAATCCCATCACCATCTT-3' and Reverse 5'- CGCCCCACTTGGATTGGG-3'

Protein Extraction And Western Blotting

Total cell lysates were isolated using cell lysis buffer (RIPA buffer: Cell Signaling Technology #9806, Danvers, MA) containing proteinase inhibitor cocktail (Roche, Nutley, NJ). Protein concentrations were determined by BCA assay (Sigma-Aldrich, St. Louis, MO). Proteins were separated by SDS-PAGE and transferred from gels to 0.2 µm nitrocellulose membranes (Pall Corporation, Washington, NY). The nitrocellulose membrane was further incubated overnight at 4°C with rabbit anti-VGLL3 (1:1000, Novus Biologicals, NBP2-31590, Centennial, USA) and rabbit anti-vinculin (1:1000, Cell Signaling Technology, 4650S, Danvers, MA). After undergoing binding with a HRP-conjugated anti-Rabbit IgG (1:1000, Cell Signaling Technology, 7074S, Danvers, MA) secondary antibody, protein bands were visualised using western blotting luminol reagent (Santa Cruz Biotechnology, Inc., Dallas, Texas).

Patients’ Tissue Samples And Clinical Information

The unstained slides from 84 HGSOC and 66 adjacent normal ovarian epithelial TMA blocks and their corresponding sets of clinical information were obtained from the Korea Gynaecologic Cancer Bank (KGCB) of Gangnam Severance Hospital Yonsei University College of Medicine (No. HTB-P2021-5). The KGCB is a part of the Bio & Medical Technology Development Program of the Ministry of the National Research Foundation (NRF), which is funded by the Korean Government Ministry of Science and ICT (MSIT) (NRF-2017M3A9B8069610). All the patients were treated with first line chemotherapy. Tissue samples and medical records were obtained with the approval of the Institutional Review Board of Gangnam Severance Hospital (IRB#, HTB-P2021-5), Seoul, Republic of Korea. The whole study was approved by the Institutional Review Board as well as the Institutional Research Medical Ethics Committee of Samsung Medical Center (IRB#, SMC 2021-06-23), Seoul, Republic of Korea. All procedures
were conducted in accordance with the guidelines of the Declaration of Helsinki. Tumour staging was performed according to the classification established by the International Federation of Gynaecology and Obstetrics (FIGO). For all study participants, CA125 levels were measured at primary diagnosis up to 1 week preoperatively using Elecsys CA125 II ECLIA (Roche Diagnostics, Rotkreuz, Switzerland). The demographics and clinical characteristics of the individuals that participated in this study are listed in Table 2.

Table 2
Comparison of VGLL3 protein expression in high-grade serous ovarian carcinoma and normal adjacent tissues

<table>
<thead>
<tr>
<th>Variables</th>
<th>No (%)</th>
<th>VgLL3_Nucleus</th>
<th>VgLL3_Cytoplasm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean IHC score (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Normal</td>
<td>66</td>
<td>41.07 [27.69–54.45]</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>HGSOC</td>
<td>84</td>
<td>133.28 [121.42–145.14]</td>
<td></td>
</tr>
</tbody>
</table>

Tumour Infiltrating Immune Cell Estimation

The estimated abundance of tumour-infiltrating immune cells (TIICs) was calculated using immunodeconv in R, which was downloaded from the Tumour Immune Estimation Resource (TIMER) 2.0 website (http://timer.cistrome.org/) [39]. To elaborate, TIICs were inferred using three different tools, including EPIC (http://epic.gfellerlab.org/), TIMER (https://cistrome.shinyapps.io/timer/), and CIBERSORT (https://cibersort.stanford.edu/). These three computational algorithms use deconvolution-based approaches that model gene expressions as the weighted sum of the expression profiles of the admixed cell types. CIBERSORT expresses results as a fraction relative to the total immune cell content. CIBERSORT abs mode transforms the results into a score in terms of arbitrary units that reflect the absolute proportion of each cell type. On the other hand, EPIC offers results as a cell fraction relative to all cells in the sample, while TIMER produces a score in terms of arbitrary units that are comparable between samples [40–42]. The outlier of TIICs was eliminated using Tukey's method. Then, a correlation analysis between the abundance of TIICs and gene expressions was conducted using Pearson's method. A p-value < 0.05 and a correlation co-efficient R ≥ 0.30 were considered to represent a significant correlation.

Immunohistochemical Analysis

To assess the clinical significance of VGLL3 protein in patients with HGSOC, immunohistochemical analysis with human ovary TMAs was performed [43]. For immunohistochemical staining, deparaffinised and rehydrated sections were retrieved via microwave for 10 min in a pH 6.0 citrate buffer. Endogenous
peroxidase in the sections was then inactivated using a peroxidase blocking solution (Agilent, S2023, Dako, Glostrup, Denmark) for 20 minutes. Next, the tissue samples were incubated with the anti-VGLL3 primary antibody (1:200, Novus Biologicals, NBP2-31590, Centennial, USA) for 2 h at 25°C. The secondary antibody was applied for 1 h at 25°C, after which detection was performed using DAB Substrate-Chromogen solution (Agilent, K5007, Dako, Glostrup, Denmark). Lastly, the sections were counterstained using haematoxylin and mounted.

Stained TMA slides were digitalised using the NanoZomer XR digital pathology (NDP) system (Hamamatsu, Hamamatsu City, Japan) at ×40 objective magnification with a single-focus layer. Digitalised images were automatically analysed using Visiopharm software version 6.9.1 (Visiopharm, Hørsholm, Denmark). Blue-coloured (haematoxylin) tumour nuclei were initially defined, after which brown-coloured (DAB) nuclei and cytoplasm were spectrally separated. Regarding the expression value of VGLL3 nuclear staining, a brown nuclear staining intensity (0 = negative, 1 = weak, 2 = moderate, and 3 = strong) and a respective percentage were obtained. For VGLL3 cytoplasmic assessment, a brown cytoplasmic intensity (weak and strong) was obtained, and each proportion was analysed. Histoscores were calculated by multiplying the percentage of positive cells by their staining intensity.

**Differentially Expressed Gene Analysis**

Data on raw read counts downloaded from GDAC were used as inputs. Samples were considered as being indicative of either high or low VGLL3 mRNA expression. Differentially expressed genes (DEGs) were identified using the DESeq2 package in R, and the cut-off value of FDR (offered as adjusted \( p \)-value) was 0.001 (\( p \)-value < 2.04e-4) [44]. Using the identified DEGs correlated with VGLL3, pathway analysis was performed using over-representation analysis based on the 19,990 predefined gene sets of Consensus PathDB ([http://cpdb.molgen.mpg.de/](http://cpdb.molgen.mpg.de/)) [45]. Next, among the identified pathways (\( p \) value < 0.01, \( q \) value < 0.2), cancer-related pathways were selected through a literature review and manual curation. We also conducted a heatmap analysis to identify the significant genes that were associated with selected pathways and high VGLL3 expression. The Complex heatmap package in R was used for visualization. Genes that were part of both the gene list from the identified pathways and the cancer gene lists of the Bushman Laboratory were selected as input. The Bushman Laboratory cancer gene list consists of 2,579 genes that are known to be human cancer genes and human homologs of cancer genes in model organisms.

**Statistical Analysis**

Data were statistically analysed using R software version 4.0.2. (R 4.0.2, Auckland, New Zealand). All mean values are expressed as mean ± standard error of the mean (S.E.M). To compare gene expression among groups with different clinical and pathological features, the DESeq2 package, Mann-Whitney test, and Kruskal-Wallis test were used. For survival analysis, Kaplan-Meier plot and log-rank test were conducted using the survival and survminer packages in R [46, 47]. To identify the independent
prognostic factor, Cox regression analyses were performed and visualised using the forest plot package in R [48]. $p$-values < 0.05 were considered to be statistically significant.

**Results**

**Patients’ characteristics**

HGSOC tumours gathered from 302 TCGA and 84 TMA datasets were analysed. The clinic pathological characteristics are described in Table 1. TMA data was further analysed based on age, pre-CA125 level, and chemosensitivity after initial treatment. CA125 levels of > 35 IU/mL ($n = 77$) were considered to be positive, while levels of $\leq 35$ IU/mL ($n = 7$) were considered to be negative.

**Altered VGLL3 mRNA expression had a prognostic significance in HGSOC**

First, we analysed the expression levels of six key genes in the hippo pathway ($YAP1$, $TAZ$, $TEAD3$, $TEAD4$, $VGLL3$, and $VGLL4$) in HGSOC. We found that the expressions of $VGLL4$, $TEAD3$, $TEAD4$, and $YAP1$ were all increased in HGSOC. By contrast, the expressions of $VGLL3$ and $TAZ$ were low in HGSOC (Fig. 1A). The result was comparable to those obtained using datasets that included the additional GSE9891 cohort (Supplementary Fig. 1A). Interestingly, $VGLL3$ expression was the lowest among all the genes in both the TCGA ($p = <2e^{-16}$) and GSE9891 cohorts ($p = <2e^{-16}$) (Fig. 1A S1A). When tumours were compared to normal ovarian samples, we found that $VGLL3$ was significantly lower in tumour samples than it was in normal ovarian samples in both the TCGA and GSE26712 cohorts ($p < 0.05$ and $p = 2.4e^{-07}$, respectively) (Supplementary Fig. 1B, C). On the other hand, $TEAD4$ was found to be significantly higher in tumour samples than it was in normal ovarian samples in both the TCGA and GSE26712 cohorts ($p < 0.05$ and $p = 5.8e^{-07}$) (Supplementary Fig. 1B, C). There were no significant differences observed regarding the expressions of $YAP1$ and $VGLL4$ between tumour and normal ovaries in both the TCGA and GSE26712 cohorts. However, the expression of $TAZ$ was discrepant between different datasets; it was significantly lower in HGSOC than it was in the normal ovary in the TCGA cohort ($p < 0.05$), but it was significantly higher in the GSE26712 data ($p = 2e^{-06}$) (Supplementary Fig. 1B, C). Finally, no significant difference was observed in $TEAD3$ expression in the TCGA cohort, whereas in the GSE26712 cohort, $TEAD3$ was significantly increased among tumour tissues compared to in normal ovary, $p = 2.6e^{-05}$. The low expression of $VGLL3$ in OC was found to be uniquely consistent in both the TCGA and GSE datasets.

Finally, we checked the expression of $VGLL3$ in different OC cell lines and found the expression of $VGLL3$ mRNA was lower in YDOV-151, YDOV-157, YDOV-13, YDOV-139, and OVCA-433 cell lines but higher in OVCAR3, YDOV-161, SKOV3 and OVCA-429 cell lines (Supplementary Fig. 1D). On the other hand, $VGLL3$ protein expression was lower in SKOV3, OVCAR3, OCVA433, YDOV-151 and YDOV-157 cell lines but higher in OVCA-429, YDOV-13, YDOV-139 and YDOV-161 cell lines (Supplementary Fig. 1E).

Next, we explored the correlation of the above six hippo-related genes with the progression of HGSOC. $VGLL3$ expression was found to be significantly increased in the advanced stages (stage III + IV) of HGSOC tumours ($p = 0.046$) compared to in the early stage II (Fig. 1B). However, $VGLL4$ ($p = 0.34$), $TEAD3$
(p = 0.093), TEAD4 (p = 0.991), YAP1 (p = 0.102), and TAZ (p = 0.05) did not show any significant difference between stage II vs. stage III + IV in HGSOC (Fig. 1B). When we checked the expressions of those six genes in terms of different grades of HGSOC, we did not observe any significant differences in any genes between (G1 + G2) vs. (G3 + G4). However, for VGLL3 expression, there was a decreasing trend observed in (G3 + G4) relative to (G1 + G2) (p = 0.098) (Fig. 1C).

Next, we checked the correlations of those six genes with the OS of patients with HGSOC. The mRNA expression of each gene was categorised into high and low groups (Supplementary Fig. 2). The data showed that only VGLL3 had a significant association with OS (p = 0.003) (Fig. 2), and this finding was consistent in both the GSE26712 & GSE9891 (p = 0.048, p = 0.012, respectively) cohorts (Supplementary Fig. 2B, C right panel). VGLL3\textsuperscript{high} correlated with the lower OS in HGSOC, while VGLL3\textsuperscript{low} was associated with better OS. However, the other five genes did not show any significant correlation with OS (Fig. 2). Taken together, these findings suggest that the role of VGLL3 mRNA is distinctive in HGSOC compared to the other five genes, which correlated with the characteristics of advanced tumour stages of HGSOC and poor survival outcomes.

VGLL3 mRNA expression was associated with macrophage infiltration and unfavourable prognosis

HGSOC is a highly heterogeneous type of cancer, and its prognosis is affected by various factors such as the tumour microenvironment and immune cell infiltration. We checked the correlation of our target genes with immune cell infiltration. Our analysis revealed that VGLL3 was positively correlated with infiltrations of macrophage, CD4 + T cell, and CD8 + T cell in HGSOC, while it was negatively correlated with B cells (Fig. 3A–C). Interestingly, the correlation of VGLL3 with macrophage infiltration was strongest among all other TIICs in HGSOC. This finding was consistently revealed in all three computational tools (EPIC, R = 0.38; p = 1.8e⁻¹¹, TIMER, R = 0.44; p = 2.7e⁻¹⁵, and CIBERSORT, R = 0.44; p = 7.2e⁻¹⁶). However, we could not find any significant correlation between the other five genes and HGSOC TIICs (Supplementary Fig. 3). Then, we checked the correlation of VGLL3 with two different subtypes of macrophage (M1 and M2 macrophage). We noted that, while there was no significant correlation found between VGLL3 with M1 macrophage (Supplementary Fig. 5A), VGLL3 was found to be significantly correlated with M2 macrophage infiltration R = 0.44; p = 5.7e⁻¹⁶ (Supplementary Fig. 5B), thus suggesting that the association of VGLL3 with macrophage mainly comes from M2 macrophage.

Next, we investigated the association of macrophage infiltration with OS in HGSOC. We observed a similar trend to that that has been observed in VGLL3, in which high levels of macrophages were significantly correlated with poor overall survival (p = 0.0057), whereas low level of macrophage expression correlated with better OS (Fig. 4A). Multivariate analysis further confirmed that VGLL3 and macrophages were independent prognostic factors for HGSOC (p = 0.003 and p = 0.024, respectively) (Fig. 4B and S4), thus suggesting that VGLL3 serves as an independent unfavourable prognostic marker in HGSOC, possibly in association with macrophage infiltration.

VGLL3 protein expression had similar effects on clinical outcomes as the VGLL3 mRNA expression
Next, to investigate the significance of VGLL3 protein expression in HGSOC, we performed immunohistochemistry-guided TMA using HGSOC tissues. The immunohistochemistry scores were analysed to check the expression of VGLL3 protein. During the analysis, we dichotomised the expression of VGLL3 into nuclear and cytoplasmic levels. A non-parametric Mann-Whitney U test was conducted to compare HGSOC vs. adjacent normal ovary. Our data showed that VGLL3 expression at both nuclear and cytoplasmic levels was significantly higher ($p < 0.001$) in HGSOC tissues than it was in the normal adjacent tissues (Table 2 and Fig. 5A, B). Then, we investigated the correlation of VGLL3 protein expression in different clinicopathological features of HGSOC and conducted the non-parametric Mann-Whitney U test (Table 3). There was no significant difference found regarding the expression of VGLL3 in different groups including age ($\leq 50$ years' vs. $>50$ years), FIGO stage (I + II vs. III + IV), tumour grade (G3 vs. G1 + G2), Pre-CA125 level (high vs. low), and chemo-sensitivity (sensitive vs. resistant) (Table 3). To evaluate the prognostic role of VGLL3 protein in HGSOC, we applied Kaplan–Meier survival analysis by determining the VGLL3$^{\text{high}}$ and VGLL3$^{\text{low}}$ groups in a manner similar to that of the analysis of the mRNA data. We found that high expressions of VGLL3 protein at both nuclear and cytoplasmic levels were correlated with lower OS in patients with HGSOC, although these correlations were not statistically significant (Fig. 5C, D).
Table 3
VGLL3 protein expression in high-grade serous ovarian carcinoma according to the clinicopathological characteristics

<table>
<thead>
<tr>
<th>Variables</th>
<th>No (%)</th>
<th>VgLL3_Nucleus</th>
<th>VgLL3_Cytoplasm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean IHC score (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>35  42.0</td>
<td>125.55 [108.31-142.79]</td>
<td>0.336</td>
</tr>
<tr>
<td>&gt;50</td>
<td>49  58.3</td>
<td>138.8 [124.23-153.37]</td>
<td></td>
</tr>
<tr>
<td><strong>FIGO stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I/II</td>
<td>13  15.5</td>
<td>153.65 [125.48-181.82]</td>
<td>0.148</td>
</tr>
<tr>
<td>III/IV</td>
<td>71  84.5</td>
<td>128.97 [116.83-141.11]</td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1 + G2</td>
<td>40  47.6</td>
<td>126.92 [110.78-143.07]</td>
<td>0.250</td>
</tr>
<tr>
<td>G3</td>
<td>44  52.4</td>
<td>139.05 [123.66-154.45]</td>
<td></td>
</tr>
<tr>
<td><strong>Pre CA-125 level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative (≤ 35U/ml)</td>
<td>7  8.3</td>
<td>146.08 [107.32-184.83]</td>
<td>0.487</td>
</tr>
<tr>
<td>Positive (&gt;35U/ml)</td>
<td>77 91.7</td>
<td>132.11 [120.43-143.80]</td>
<td></td>
</tr>
<tr>
<td><strong>Chemosensitivity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitive</td>
<td>67  79.8</td>
<td>137.60 [125.19–150.00]</td>
<td>0.270</td>
</tr>
<tr>
<td>Resistant</td>
<td>15  17.9</td>
<td>122.17 [95.95-148.38]</td>
<td></td>
</tr>
</tbody>
</table>

A non-parametric Mann-Whitney U test was performed to compare each diagnosis parameter.

DEGs and pathway analysis identified altered pathways associated with VGLL3

We investigated VGLL3-related pathways in cancer to explore the potential mechanism in HGSOC. The DEG analysis showed that around 3,981 genes were significantly associated with VGLL3 expression.
Pathway analysis using 3,981 DEGs revealed that the gene sets associated with VGLL3 mRNA expression showed enhancements in extracellular matrix (ECM) organization, focal adhesion, PI3K-Akt signalling pathway, and JAK-STAT signalling pathway (Fig. 6A). The association of VGLL3 with those pathways has previously been reported in different cancers [26]. Interestingly, along with previously reported pathways, we identified three novel pathways that were associated with VGLL3: Nonsense-Mediated Decay (NMD), vascular endothelial growth factors A-vascular endothelial growth factor receptor 2 (VEGFA-VEGFR2) signalling pathway, and platelet-derived growth factor (PDGF) signalling pathway (Fig. 6B). For further investigation into the molecular mechanism of VGLL3, a heatmap of DEGs was created between VGLL3high and VGLL3low (Fig. 7). We showed that those genes had a strong correlation with VGLL3 expression based on the z scores. We observed a distinguishable positive vs negative correlation between those genes with high vs. low VGLL3 mRNA expressions (Fig. 7).

**Discussion**

In recent years, the role played by VGLL3 in cancers has attracted increasing research attention because of its dynamic behaviour in different cancers. VGLL3 has been found to play dual roles in cancers [23–27]. In this study, we found that VGLL3 was an independent unfavourable prognostic factor of HGSOC, likely by affecting immune cell infiltration, particularly macrophages, and regulating or deregulating a significant number of oncogenic genes and pathways. The results suggest that VGLL3 is an important factor for HGSOC that correlates with tumour progression and immune evasion.

The malignancy of EOC largely depends on the constitutive activation/deactivation of different oncogenes, tumour suppressor genes, and transcription factors [49, 50]. Dysregulation of these genes is associated with tumour progression, recurrence, and metastasis. The correlation of VGLL3 in cancer proliferation, advanced tumour stage, grade, and poor prognosis has previously been reported in other cancers except for ovarian cancer [25, 26]. In line with previous reports, we found that VGLL3 expression in both mRNA and protein level was also correlated with advanced tumour stage and poor prognosis in HGSOC, suggesting that VGLL3 may promote the progression of HGSOC. Importantly, we may suggest a distinctive role of VGLL3 that conflicts with the previous report by Gambaro et al. where VGLL3 induction showed a tumour-suppressive phenotype in EOC [27]. We dispute their hypothesis based on their contradictory experimental results that VGLL3 was undetectable in the parental OV-90 cell line, and was only observed after chromosome transfer, as a single gene failed to suppress cell proliferation and tumour growth both *in vitro* and *in vivo* and based on our findings.

In contrast to the reduced expression of VGLL3 mRNA in HGSOC, we found that VGLL3 protein expression was increased in HGSOC. We also observed this opposite expression pattern in some our OC cell line study especially SKOV3, OVCAR3 and YDOV139 cell line. A decent explanation for this discrepancy is currently unknown; one possible hypothesis might involve the effect of post-transcriptional modification. For example, mRNA regulatory elements and the affinity of RNA Binding Proteins (RBPs) are linked to increasing RNA stability and translational efficiency of mRNA molecules, which leads to the aberrant expression of protein in tumour cells [51, 52]. Therefore, we assume that post-transcriptional
modications, such as of VGLL3 mRNA regulatory element and RBD, are more active in HGSOC than in the typical case, which may increase VGLL3 mRNA stability as well as protein translation, and thereby ultimately increase its overexpression in HGSOC. Alternatively, the post-translation modification of VGLL3 protein might also play a role in stabilizing the VGLL3 protein, preventing their degradation, and thus increasing VGLL3 protein in HGSOC tumours. There is a need for an in-depth study to explore the post-transcriptional regulation of VGLL3 and its effect on VGLL3 mRNA stability and protein expression. Despite the fact that there is a discrepancy in VGLL3 mRNA and protein expression between tumour and normal tissues, we interestingly observed that higher levels of both VGLL3 mRNA and protein expression among tumour tissues were associated with the worse OS in HGSOC.

Tumour-infiltrating lymphocytes and the immune status of the tumour microenvironment have been reported to affect progression, therapeutic effects, and recurrence in many cancers [53]. Macrophages in the tumour microenvironment play vital physiological and pathological functions. Tumour-associated macrophages (TAMs), which mainly belong to the M2 macrophage phenotype, are known to correlate with poor outcomes in solid cancers and play important roles in tumourigenesis [54]. A high density of CD163+ M2-macrophages is predominantly associated with poor prognosis in ovarian cancer and known to be involved in tumour invasion, angiogenesis, metastasis, and early recurrence [55, 56]. Moreover, the high M1/M2 ratio of tumour infiltrating macrophages correlates with prolonged survival time in EOC, while a low M1/M2 ratio correlates with poor OS [57, 58]. In this study, we also observed the association of VGLL3 with macrophage infiltration in HGSOC, which likely contributed to the worsening of OS. Moreover, we found that M2 macrophage was more strongly correlated with VGLL3 than M1 macrophage, suggesting that VGLL3 may be involved in the poor prognosis of HGSOC by association with macrophages, particularly M2 macrophage.

In addition to known clinical and molecular biomarkers such as TP53, BRCA1/2, and MYC, VGLL3 regulates many key genes and pathways, and it is also related to clinical prognosis. High VGLL3 expression has been found to activate several signalling pathways, such as MAPK, JAK-STAT, PI3K/Akt/mTOR, ECM, focal adhesion, and WNT pathways in tumours [25, 26]. In our study, we also discovered the pathways in HGSOC that correlated with high VGLL3 expression. Further, we found three novel pathways (NMD, VEGFA-VEGFR2, PDGF) that were associated with high VGLL3, suggesting that VGLL3 may regulate key genes in those pathways. The frequent activation of ECM, PI3K/Akt/mTOR, and VEGFA-VEGFR2 signalling pathways has been associated with higher invasive and migratory capacities in subpopulations of human OC [59–61]. TAMs have been reported to correlate with many signalling pathways including PI3K/AKT/mTOR signalling pathway, ECM, and focal adhesion molecules which modulate the tumour microenvironment [62–64]. In this report, we connected VGLL3 mRNA expression with macrophage infiltration and multiple signalling pathways; we therefore speculate that VGLL3 may play a critical role in the poor prognosis of HGSOC by activating TAMs and molecules related to signalling pathways.

Obviously, the limitation of our study is that it used public mRNA data not derived from our own patients and lacked evidence from in vitro and in vivo research. Nonetheless, the public DBs including TCGA
contain decent numbers of curated data, and the findings were highly correlated with our immunohistochemistry-guided TMA in HGSOC tissue samples. Therefore, VGLL3 has been strongly suggested to have significance and a potential role in both mRNA and protein levels, which represents the strength of this study. Currently, we are conducting research into investigating the molecular mechanism of VGLL3 in HGSOC as a continuation of this study.

**Abbreviations**


**Conclusion**

In this study, we can conclude that VGLL3 has potential prognostic values in HGSOC because its overexpression was shown to be associated with advanced tumour stage, TAM infiltration, and poor prognosis. The DEGs pathways and co-expressing genes identified in this study provide further insights into the molecular basis of the role of VGLL3 in HGSOC. There is still a need for further research to show the direct interaction and functional interplay of VGLL3 with related molecules. Finally, our study of VGLL3 and its association with macrophage infiltration will help us better understand the immune suppressive microenvironment of cancer, which will eventually guide us to identify new therapeutic solutions to treat cancers.

**Declarations**

**Ethics approval and consent to participate**

Tissue samples and medical records were obtained with the approval of the Institutional Review Board of Samsung Medical Center (IRB#, SMC 2021-06-23), Seoul, Republic of Korea, and Institutional Review Board of Gangnam Severance Hospital (IRB#, HTB-P2021-5), Seoul, Republic of Korea. All procedures were conducted in accordance with the guidelines of the Declaration of Helsinki. All participants in this study provided written informed consent.

**Consent for publication**

Not applicable

**Availability of data and materials**

All data generated or analysed during this study are included in the published article. Further
inquiries can be directed to the corresponding author.

**Competing interest**


**Funding**

This study was supported by the National Research Foundation of Korea (NRF-2019R1A2C2088715), which is funded by the Korean government (The Ministry of Science and ICT), South Korea.

**Author contributions**

Haque R, Yun JW and Kang ES contributed to the conception and design of the study. Haque R, Lee J, and Yun JW curated the data from TCGA and GEO databases and performed analysis. Kim JH and Shin HY collected the tissue samples. Chung JY, Shin HY, and Kim JH performed TMA guided immunohistochemistry and analysis. Kim H and Kim JH curated the clinical information. Haque R, Lee J, Yun JW, and Kang ES performed the statistical analysis and interpretations. Haque R, Lee J, Yun JW, and Kang ES wrote the original draft of the manuscript. Kang ES supervised the whole study and were involved in funding acquisition. All authors carefully read, edited and approved the final manuscript.

**Acknowledgement**

Not applicable

**References**


**Figures**

![Figure 1](image_url)

**Figure 1**

Correlation between mRNA expression and clinicopathological features in The Cancer Genome Atlas (TCGA) ovarian cancer data. Samples without RPKM and Raw read counts data were omitted. (A) Levels of mRNA expression among six genes. RPKM data from 296 samples were used, and the p-values were calculated using the Kruskal-Wallis test. (B) Levels of mRNA expression among different stages of HGSOC: Stage I (n = 0), Stage II (n = 18), Stage III (n = 239), Stage IV (n = 36). The p-value was calculated...
using the DESeq2 package. (C) Levels of mRNA expression among different grades of HGSOC: G1 (n = 1), G2 (n = 33), G3 (n = 254), G4 (n = 1). The p-value was calculated using the DESeq2 package. Each point represents an individual sample.

**Figure 2**

Kaplan–Meier curves for overall survival of six target genes. Normalised read counts data for 302 The Cancer Genome Atlas (TCGA) ovarian cancer samples were analyzed for overall survival while comparing the high and low expression of each gene; Top: from left to right VGLL3 (p=0.003), VGLL4 (p=0.6), TEAD3 (p=0.48); Bottom: from left to right TEAD4 (p=0.38), YAP1 (p=0.81), TAZ (p=0.44). High and low expression of each gene was selected based on the visual distinction described in Supplementary Fig. 2. Statistical significance was evaluated using the log-rank test.
Figure 3

Correlation between \textit{VGLL3} mRNA expression and immune cell infiltration. The levels of immune cell infiltrations were estimated using three databases: (A) EPIC, (B) TIMER, and (C) CIBERSORT abs mode. The three methods show that macrophage and \textit{VGLL3} reached a consensus on a significantly positive correlation. The X-axis is the estimated values of three algorithms that represent immune cell fractions and the Y-axis represents the \textit{VGLL3} mRNA expression. For CIBERSORT, cell fractions for each immune cell were considered as a summation of their subsets. After eliminating outliers of immune cell fractions using Tukey's method, Pearson's method was performed to determine the correlation between \textit{VGLL3} gene and the immune cells (\textit{from left to right}: B cell, CD4+ T cell, CD8+ T cell and macrophage), and the correlation coefficient was shown as R. For 303 The Cancer Genome Atlas (TCGA) ovarian cancer samples, the normalised read counts from TCGA Genome Data Analysis Center (GDAC) database and the immune cell infiltration levels from TIMER 2.0 website were downloaded and used to draw plots. A $p < 0.05$ and $R \geq 0.3$ was considered as statistically significant.
Figure 4

*VGLL3* and macrophages are unfavorable prognostic markers. Normalised read counts and infiltrated immune cell fractions estimated using EPIC for 303 HGSOC tissue samples were utilised. (A) Kaplan–Meier survival curves of OS comparing high (n=44) and low (n=259) macrophage infiltration in HGSOC, \( p = 0.0057 \). (B) Forest plot visualizing the hazard ratio with a 95% confidence interval and \( p \)-value was calculated using multivariate Cox regression analysis. The levels of infiltrated immune cells estimated using EPIC were multiplied by 100 to transform them into percentile values. All variables were considered to be continuous variables.
Expression of VGLL3 protein in normal ovary and HGSOC tissues and survival analysis depending on the expression levels of VGLL3 on nucleus and cytoplasm of cells. The unstained TMA blocks from 84 HGSOC and 66 adjacent normal ovarian tissues were immunohistochemically stained with anti-VGLL3 antibody. (A) Representative image of VGLL3 expression between HGSOC and adjacent normal ovary tissues. (B) IHC scores of VGLL3 in both nucleus and cytoplasm were calculated and compared between normal ovary and HGSOC tissues. (C) Survival analysis in the nucleus showing a trend towards a difference in survival probability between high and low expression, with a p-value of 0.097. (D) Survival analysis in the cytoplasm showing no significant difference in survival probability between high and low expression, with a p-value of 0.75.
HGSOC and adjacent normal ovary, p <0.0001. (C, D) Kaplan–Meier survival curves of OS comparing high (n=18) and low (n=66) expression of VGLL3 protein in HGSOC are shown both in nucleus (p=0.097) and cytoplasm (p=0.75). Each point represents an individual sample.

Figure 6

Pathway analysis using over-representation analysis from Consensus PathDB (CPDB). Overall, 3,981 differentially expressed genes (DEGs) were identified after false discovery rate control (adjust $p$-value < 0.001, $p$-value < 2.04e-4). The cutoff $p$-value and $q$-value were 0.01 and 0.2 in over-representation analysis, respectively. Four previously reported pathways and three novel pathways were captured. Raw read counts for 296 samples from The Cancer Genome Atlas (TCGA) Genome Data Analysis Center (GDAC) database were used for DEGs analysis; *, The number of DEGs overlapped in the pathway / The total number of genes of the pathway.
Figure 7

Heatmap of DEGs between the $VGLL3^{\text{high}}$ and $VGLL3^{\text{low}}$ group. For the heatmap, the z score normalised value from the results of DESeq2 were used. The genes presented in the heatmap were selected genes that were overlapped with the cancer gene list provided by Bushman laboratory (See Methods).Previously reported pathways had 128 overlapped genes in total, and novel pathways had 55 overlapped genes in total. * previously reported pathways; ** newly found pathways in this study.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.tif
- FigureS2.tif
- FigureS3.tif
- FigureS4.tif
- FigureS5.tif
- Graphicalabstract.png