Original Article

Signature of arylacetamide deacetylase expression is associated with prognosis and immune infiltration in ovarian cancer

Jianyang Feng, DEGREE, Hong He, DEGREE

Department of Gynaecology and Obstetrics, the Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

Running title: Prognostic AADAC signature in ovarian cancer

Corresponding author:

Hong He, DEGREE

Department of Gynaecology and Obstetrics, the Third Affiliated Hospital of Guangzhou Medical University, ADDRESS, Guangzhou 510000, China

E-mail: hehe2000010@gzhmu.edu.cn

https://orcid.org/0000-0002-5335-5433

Received: 2021.07.20.
Revised: 2021.10.18.
Accepted: 2021.11.02.
**Objective:** The role of the protein-coding gene arylacetamide deacetylase (*AADAC*) in the prognostication of ovarian cancer remains uncertain. We aimed to identify and validate its prognostic value using integrated bioinformatics analyses.

**Methods:** Gene expression profiles of RNA-sequencing and microarray data were retrieved from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO). Univariate and multivariate Cox regression models were used to evaluate the prognostic value of gene expression. The predictive accuracy of the gene signature model was evaluated using a time-dependent receiver operating characteristic (ROC) curve. In addition, the correlation between immune infiltration and *AADAC* was identified. A nomogram of the gene signature with clinical parameters was constructed to estimate the clinical application of the signature for survival prediction in patients with ovarian cancer.

**Results:** Univariate and multivariate Cox regression analyses in the training and validation cohorts indicated that a high *AADAC* expression signature was significantly and independently correlated with better survival outcomes in ovarian cancer. *AADAC* upregulation positively correlated with the infiltration of CD4+ memory T cells. Immunological signature gene sets were significantly enriched in CD4+ T cell regulation pathways. The area under the curve of the time-dependent ROC for overall survival indicated that the constructed nomogram had a moderate predictive ability for prognostic prediction in ovarian cancer.

**Conclusion:** *AADAC* expression signature significantly and independently correlated with the survival outcome and CD4+ memory T cell infiltration in ovarian cancer, indicating its potential applicability in the prediction of prognosis and immunotherapy efficacy.

**Keywords:** Arylacetamide deacetylase; Ovarian cancer; Bioinformatics analysis; Prognosis; Immune infiltration; Nomogram
Introduction

Ovarian cancer is one of the most fatal malignancies and the leading cause of mortality in cancers associated with the female genital tract, with an estimated 184,799 deaths reported worldwide in 2018 [1]. More than 70% of ovarian cancer cases are diagnosed at an advanced stage and are accompanied by disseminated peritoneal diseases [2]. Distinct molecular phenotypes in ovarian cancer show heterogeneous treatment responses and biological behaviors. For instance, although the high-grade serous subtype is more sensitive to chemotherapy, it is also more aggressive than low-grade subtypes. Thus, identification of gene expression signatures with key clinicopathological parameters may facilitate the precise estimation of survival, and consequently lead to the development of personalized treatment strategies. Based on recent advances in genomic sequencing techniques, it is possible to explore the significant genetic variants and expression characteristics of ovarian cancer.

Arylacetamide deacetylase (AADAC) is a protein-coding gene located in human chromosome 3. The translated product of AADAC competes with cytosolic arylamine N-acetyltransferase activity in the cytochrome P450 pathway and is associated with the biotransformation, hydrolase activity, triglyceride lipase activity, and hydrolysis of clinical drugs [3-5]. AADAC is a tissue-specific protein-coding gene that is mainly expressed in normal liver tissues and shows extremely low expression in normal ovarian tissues [6]. However, the role of AADAC in the prognostication of ovarian cancer remains inconclusive. Thus, our study aimed to identify and validate the correlation between AADAC expression and survival in ovarian cancer cases using integrated bioinformatics analyses. The results of the present study indicated that ovarian cancer cases with high AADAC expression signatures were associated with better overall survival. Furthermore, upregulation of AADAC was correlated with increased CD4+ memory T cell infiltration in the tumor microenvironment. Finally, by integrating the signature score of AADAC expression with clinicopathological parameters, we constructed a nomogram to demonstrate its potential utility in predicting the survival of patients with ovarian cancer.
Materials and methods

1. Acquisition of multi-omics data on ovarian cancer

Gene expression profiles (RNA-sequencing data) of The Cancer Genome Atlas-ovarian cancer (TCGA-OV) cohort were retrieved from TCGA using the University of California, Santa Cruz, Xena browser (USCS Xena) datahub (https://xenabrowser.net/; accessed on Oct 12, 2020), which is a high-performance visualization and analysis tool for both large public repositories and private datasets [7]. The data included HTSeq-counts and fragments per kilobase of transcript per million values of 379 cases with patient follow-up information. TCGA-OV cohort was used as the training dataset.

Raw microarray data of the GSE18520 and GSE26712 datasets were accessed from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/). The GSE18520 dataset comprised of gene expression profiles from 53 high-grade primary papillary serous ovarian adenocarcinoma specimens and 10 normal ovarian surface epithelium brushing specimens, which were evaluated using the Affymetrix Human Genome U133 Plus 2.0 Array (COMPANY, CITY, STATE, COUNTRY); while the GSE26712 dataset comprised of 185 primary ovarian tumour samples and 10 normal ovarian surface epithelium samples, which were detected using the Affymetrix Human U133A microarray (COMPANY, CITY, STATE, COUNTRY). For the validation studies, we used the GSE140082 dataset, which comprised of data on the gene expression arrays from 380 formalin-fixed paraffin-embedded ovarian cancer samples and overall survival information of the patients, which was also obtained from the GEO database and evaluated using Illumina HumanHT-12 WG-DASL V4.0 R2 Expression BeadChip (COMPANY, CITY, STATE, COUNTRY). Details of the primary data used in this study are presented in Table 1.

2. Quality control and processing of the multi-omics data
The quality of the raw data obtained from the GSE18520 and GSE26712 datasets was assessed using the “arrayQualityMetrics” (version 3.44.0; COMPANY, CITY, STATE, COUNTRY) package [8]. A robust multi-array average algorithm was used for data normalization and background correction. Different microarray scanning times were considered as a known batch effect, which was examined using the ComBat function of the “sva” (version 3.36.0; COMPANY, CITY, STATE, COUNTRY) package [9]. Batch effect control was performed by using a principal component analysis through the “FaceFactoMineR” (version 2.4; COMPANY, CITY, STATE, COUNTRY) package [10]. The gene annotation file (gencode.v35.annotation.gtf.gz) of the human genome was used to extract data on protein-coding genes and was downloaded from GENCODE (https://www.gencodegenes.org/; accessed on Oct 12, 2020). Quality annotations of TCGA-OV samples were performed according to the merged_sample_quality_annotations.tsv file (https://gdc.cancer.gov/about-data/publications/pancanatlas). The workflow of the quality control process of TCGA-OV samples is shown in Supplementary Fig. 1. A variance inflation factor was used to detect the severity of multicollinearity between the gene signature and clinicopathological parameters.

3. Differential expression analysis of the protein-coding genes

Differential expression analysis of the protein-coding genes between the normal and cancerous ovarian samples obtained from the GEO datasets was performed using the “limma” (version 3.44.3; COMPANY, CITY, STATE, COUNTRY) package [11]. The batch effects of the scanning time in the microarray datasets were integrated into the limma model design to conduct further differential expression analysis. Genes with |logFC| >1 and adjusted $P<0.05$, were considered differentially expressed.

4. Identification of survival-associated protein-coding genes
The value of protein-coding gene expression on survival was determined using a univariate Cox proportional hazards model analysis (log-rank $P<0.05$) in the TCGA-OV training cohort.

5. Univariate and multivariate Cox regression model analyses and validation

The intersection data between the potential survival-associated genes and differentially expressed genes were filtered using a least absolute shrinkage and selection operator (LASSO)-penalized Cox regression analysis. Only genes with non-zero coefficients via the minimum criteria (a $\lambda$ value of 0.066 with log $\lambda=-2.718$ were selected by a 10-fold cross validation) in the LASSO regression model were chosen to further estimate the prognostic significance using univariate and multivariate Cox regression analyses. Next, Kaplan-Meier survival curves were plotted and the log-rank method was used to estimate the prognostic significance. The signature score of a specific gene was calculated by multiplying the gene expression values by the coefficients. A time-dependent receiver operating characteristic (ROC) curve was plotted to investigate the predictive accuracy using the “survivalROC” (version 1.0.3; COMPANY, CITY, STATE, COUNTRY) package [12]. Finally, the validation cohort GSE140082 was used to verify the prognostic significance of the gene signature.

6. Estimation of immune infiltration and immunological signature gene set enrichment in ovarian cancer

xCell [13], a webtool and gene signature-based method was used to perform cell type enrichment analysis based on the gene expression data of 64 immune and stromal cell types. Data on immunological signature gene sets, which represent the cellular states and perturbations within the immune system, were downloaded from the Molecular Signatures Database v7.2 (MSigDB), Broad Institute (https://www.gsea-msigdb.org/; accessed on Oct 12, 2020).
7. Construction of the nomogram for the gene signature

In the TCGA-OV cohort, the following parameters were selected to construct a nomogram using the “survival” (version 3.2-10; COMPANY, CITY, STATE, COUNTRY) [14] and “rms” (version 6.1-1; COMPANY, CITY, STATE, COUNTRY) packages: age, stage, grade, and residual disease along with the gene signature score [15]. Calibration curves were plotted to assess the concordance between the actual and predicted survival. A time-dependent ROC curve was plotted to investigate the predictive accuracy of the nomogram.

8. Statistical analysis

An adjusted $P$-value $<0.05$, with $|\log FC| > 1$, was considered significant for differential gene expression analysis. The Kaplan-Meier method and log-rank test were used to estimate the impact of the gene signature on survival. Statistical significance was set at $P<0.05$. A significant Spearman’s correlation coefficient was estimated with a $P$-value $<0.05$. All data were processed using the R software (version 4.0.2 [×64]; COMPANY, CITY, STATE, COUNTRY) platform.

Results

1. Clinicopathologic characteristics in the training and validation cohorts

The clinicopathologic characteristics (including age, stage, grade, residual disease, and/or debulking surgery) of the training and validation cohorts are summarized in Table 2.
2. Differentially expressed protein-coding genes in ovarian cancer

Normal ovarian samples and ovarian cancer samples from the GSE18520 and GSE26712 datasets were compared according to the adjusted $P$-value <0.05, and $|\log FC| > 1$. The batch effect, which was the scanning time in the two microarray profiles, was integrated into the differential expression analysis model and visualized using principal component analysis (Fig. 1D, E, respectively). In the GSE18520 dataset, 1,226 and 712 protein-coding genes were significantly upregulated and downregulated, while in the GSE26712 dataset, 552 and 765 protein-coding genes were upregulated and downregulated, respectively. Volcano plots of the differentially expressed genes are shown in Fig. 1A. Based on the data of these genes, an intersection of 502 differentially expressed (227 upregulated and 275 downregulated) protein-coding genes was observed, and these genes were selected for further analysis (Fig. 1B, C).

3. Survival-related differentially expressed protein-coding genes in ovarian cancer

The correlation between the differentially expressed protein-coding genes and survival in the TCGA-OV cohort was analyzed using a univariate Cox regression analysis ($P<0.05$). A total of 1,354 protein-coding genes were significantly associated with survival in ovarian cancer, and 32 intersecting protein-coding genes from TCGA-OV, GSE18520, and GSE26712 datasets were screened for prognostic value in ovarian cancer (Fig. 1F). Furthermore, data on these 32 genes were filtered using the LASSO-penalised Cox regression analysis (Fig. 1G, H). Moreover, seven genes, including isocitrate dehydrogenase ($IDH2$), FA complementation group I ($FANCI$), C-X-C motif chemokine receptor 4 ($CXCR4$), PRAME nuclear receptor transcriptional regulator ($PRAME$), praja ring finger ubiquitin ligase 2 ($PJA2$), enoyl-CoA delta isomerase 2 ($ECI2$), and $AADAC$, were identified for conducting further univariate and multivariate Cox regression validation analyses. Finally, univariate and multivariate Cox proportional hazard regression analyses performed using both the training and validation cohorts verified that $AADAC$ was significantly and independently correlated with prognosis in ovarian cancer (Table 3, Fig. 2A, 3B).
4. AADAC expression signature is associated with prognosis in ovarian cancer

The samples were classified according to the median AADAC expression score then grouped as high-and low-score signature groups. The Kaplan-Meier analysis revealed a significant difference in the outcome of the patients with ovarian cancer between the high- and low-score signature groups (log-rank test \( P<0.05 \); Fig. 3A, D). A better survival outcome was significantly correlated with the high-score AADAC expression signature in the subgroup analysis (Fig. 3C, F). The time-dependent area under the ROC curves (AUC) for 1-, 3-, and 5-year survival in the training and validation cohorts are illustrated in Fig. 3B, E, respectively. After combining the AADAC signature score with the clinicopathological parameters (age, stage, grade, residual disease, and/or debulking surgery) in both cohorts, multivariate Cox regression and subgroup analyses were conducted, which further indicated that the high-score AADAC expression signature was significantly and independently correlated with favorable prognosis in ovarian cancer (Fig. 3C, F). The evaluation of the variance inflation factor suggested that the AADAC expression signature showed less collinearity with the other clinical variables in the model (Supplementary Fig. 2). In order to illustrate the reliability of the signature, we subdivided the groups based on the upper and lower quantiles of the AADAC expression signature score into the following: top 25%, bottom 25%, and 25-75% subgroups, and the results still demonstrated a favorable prognosis in the top 25% high-score subgroup (Supplementary Fig. 3).

5. Immune infiltration and immunological signature enrichment in ovarian cancer

Immune infiltration in ovarian cancer was estimated using the xCell algorithm. We found that an increased extent of CD4+ memory T cell infiltration was significantly correlated with the upregulation of AADAC (\( \rho=0.18; P<0.001; 95\% \) confidence interval [CI], 0.22-0.42; Fig 4A, B), and was associated with a better survival outcome (Fig. 4C). Furthermore, the immunological signature enrichment analysis demonstrated that
upregulated AADAC expression was significantly enriched in genes that were overexpressed in the T cell receptor pathway and CD4+ T cell regulation (Fig. 4D).

6. Nomogram model for the AADAC expression signature with clinicopathologic parameters based on TCGA-OV cohort

We combined the AADAC expression signature score with feasible clinicopathologic parameters, including age, stage, grade, residual disease, and signature score (Fig. 5A). Each patient was assigned one point for each prognostic parameter: the higher the total number of points, the poorer the outcome. The calibration plots of the nomogram and an ideal model were compared; the nomogram showed moderate performance (Fig. 5B-D). Moreover, the time-dependent AUC of 1-, 3-, and 5-year overall survival demonstrated that our model had a moderate predictive ability (Fig. 5E).

Discussion

Ovarian cancer is the most fatal malignancy associated with the female genital tract [2,16]. Most patients are diagnosed at an advanced stage with early peritoneal dissemination at the primary diagnosis. Patients encounter a disastrous relapse even after undergoing primary debulking cytoreduction surgery and adjuvant platinum-based chemotherapy [16]. Exploration of the significant gene expression characteristics in cancerous samples can directly indicate the inherent molecular heterogeneity in ovarian cancer cells [17-19]. With the development and advancement of multiple “omics” sequencing techniques, the processing of considerable and complex sequencing data using comprehensive and integrated bioinformatics methodologies can help us better understand the underlying molecular characteristics of the cancer genome. In this study, we utilized bioinformatics approaches to estimate the prognostic impact of AADAC in ovarian cancer.
AADAC is a tissue-specific protein-coding gene in humans [5] that is highly expressed in the liver but not in the ovarian tissues [3,5,6]. The translated product of AADAC participates in various activities related to enzymatic bioactivation reactions [5,20]. Jiang et al. [21] demonstrated that intestinal AADAC is involved in vicagrel hydrolytic bioactivation to further influence platelet function. The platelet-lymphocyte ratio has been associated with survival outcomes in various cancer subtypes [22-24]. Recently, Toyohara et al.[25] found that elevated AADAC expression in vascular smooth muscle cells was correlated with cell proliferation, migration, and apoptosis.

In the present study, univariate and multivariate Cox regression analysis results indicated that AADAC overexpression was significantly associated with better survival in patients with ovarian cancer. Furthermore, survival analysis revealed that, compared with a low-score signature, a high-score signature of AADAC expression was significantly and independently related to better outcomes in patients with ovarian cancer. Similarly, Liu et al. [26] suggested AADAC expression (hazard ratio, 1.112; 95% CI, 1.042-1.186) as a prognostic gene signature with applicability in predicting the overall survival in gastric cancer via an integrated analysis of multiple gene expression profiles. Finally, we constructed a nomogram based on TCGA-OV cohort and clinicopathologic parameters, including the signature score of AADAC expression, age at diagnosis, tumor grade, tumor stage, and residual disease status. The nomogram constructed in the present study demonstrated a moderate potential for predicting ovarian cancer prognosis in clinical practice.

Immune infiltration is highly heterogeneous in ovarian cancer [27,28]. The T-cell population participates in the elicitation of an anti-tumor immune response and is associated with survival outcomes in patients with ovarian cancer [29,30]. Similar to the findings of a previous study [30], we observed that a lack of CD4+ memory T cell infiltration in the tumor microenvironment was significantly correlated with a worse outcome. The results indicated that tumor infiltration by the T-cell population reflected a tumor-related immune response. Thus, reversal of the immunosuppressive tumor microenvironment provides the possibility of adopting immunotherapy strategies for ovarian cancer treatment.
The limitations of our study must be acknowledged. First, our study lacked the experimental data necessary to validate the findings derived from the bioinformatics analyses. Second, the mechanisms underlying the regulation of AADAC expression, as underlined in the comprehensive molecular network, should be further explored. In summary, we identified and validated the correlation between the AADAC expression signature and survival outcome in ovarian cancer, highlighting its potential prognostic value. However, future research is warranted to validate these findings. The AADAC expression was integrated with age, tumor grade, tumor stage, and residual disease in clinical practice based on the model assessment performed herein; consequently, this might promote the stratification of survival prediction in ovarian cancer. Finally, upregulated AADAC expression was correlated with CD4+ T-cell infiltration and regulation, processes that participate in the development of adaptive anti-tumor immunity in the ovarian cancer microenvironment, thereby suggesting its potential utility in the prediction of immunotherapy efficacy in ovarian cancer.

In conclusion, the AADAC expression signature was associated with better survival outcome and demonstrated a favorable prognostic potential for ovarian cancer. Overexpressed AADAC correlated with an increased extent of CD4+ memory T cell immune infiltration in the tumor microenvironment. Thus, integration of AADAC expression with feasible clinicopathological parameters has potential clinical value in the prediction of prognosis and immunotherapy efficacy in ovarian cancer.

Conflict of interest

The authors declare that they have no competing interests.

Ethical approval

Not applicable.
Patient consent

Not applicable.

Funding information

None.

Acknowledgments

We thank Dr. Jianming Zeng (University of Macau), and all the members of his bioinformatics team, biotrainee, for generously sharing their experience and codes. The authors gratefully acknowledge contributions from the TCGA Network, Prof. Winterhoff B. and Stefan Kommoss S. (Mayo Clinic) for providing the publicly available GSE140082, Prof. Birrer MJ (Surgical Oncology Research Labs, Boston) for GSE18520 and GSE26712.

Supplementary material

Supplementary Fig. 1-3 associated with this article can be found online at https://doi.org/10.5468/ogs.21237.

References


Figure Legends

**Fig. 1.** Differentially expressed protein-coding genes in ovarian cancer. (A) Volcano plots of the differentially expressed genes. (B, C) Intersection of the upregulated and downregulated protein-coding genes in the GSE18520 and GSE26712 datasets. (D, E) Principal component analysis plots for performing the batch effect control. (F-H) LASSO-penalised Cox regression analyses for differentially expressed and survival-related protein-coding genes. DEG, differentially expressed gene; LASSO, least absolute shrinkage and selection operator; TCGA-OV, The Cancer Genome Atlas-ovarian cancer.

**Fig. 2.** Univariate and multivariate Cox regression analyses in (A) the training and (B) validation cohorts. CI, confidence interval.

**Fig. 3.** Kaplan-Meier plot and evaluation of the arylacetamide deacetylase (AADAC) signature score in ovarian cancer. (A, D) Kaplan-Meier plot of the AADAC signature score in the training and validation cohorts. (B, D) Time-dependent ROC curve of the AADAC signature score. (C, F) Multivariate Cox regression analysis of the AADAC signature score. ROC, receiver operating characteristic; CI, confidence interval; FIGO, The International Federation of Gynecology and Obstetrics. *EXPALIN. **EXPALIN. ***EXPALIN.

**Fig. 4.** Correlation of the arylacetamide deacetylase (AADAC) expression with immune infiltration and immunological signature in ovarian cancer. (A, B) An increased extent of CD4+ memory T cell infiltration was significantly correlated with the upregulation of AADAC. (C) Kaplan-Meier plot illustrating CD4+
memory T cell infiltration. (D) CD4+ T cell enrichment analysis in the immunological signature. CI, confidence interval; NES, FULL NAME.

**Fig. 5.** Nomogram model and evaluation of the arylacetamide deacetylase (*AADAC*) expression signature with clinicopathologic parameters in ovarian cancer. (A) Nomogram model of the *AADAC* expression signature with clinicopathologic parameters. (B-D) Calibration curves of the nomogram model for 1-, 3-, and 5-year survival. (E) Time-dependent ROC curves of the nomogram model for 1-, 3-, and 5-year survival. OS, overall survival; ROC, receiver operating characteristic.
Table 1. Characteristic details for the primary dataset

<table>
<thead>
<tr>
<th>Source</th>
<th>GEO accession</th>
<th>Platform</th>
<th>No. of tumor samples</th>
<th>No. of normal samples</th>
<th>Gene expression profiling</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGA</td>
<td>/</td>
<td>TCGA-OV</td>
<td>379</td>
<td>/</td>
<td>RNA-seq</td>
<td>Training cohort</td>
</tr>
<tr>
<td>GEO</td>
<td>GSE140082</td>
<td>GPL14951</td>
<td>380</td>
<td>/</td>
<td>Microarray</td>
<td>Validation cohort</td>
</tr>
<tr>
<td>GEO</td>
<td>GSE18520</td>
<td>GPL570</td>
<td>53</td>
<td>10</td>
<td>Microarray</td>
<td>DEA</td>
</tr>
<tr>
<td>GEO</td>
<td>GSE26712</td>
<td>GPL96</td>
<td>185</td>
<td>10</td>
<td>Microarray</td>
<td>DEA</td>
</tr>
</tbody>
</table>

GEO, Gene Expression Omnibus; TCGA-OV, The Cancer Genome Atlas-ovarian cancer; DEA, FULL NAME.
**Table 2.** Clinicopathologic characteristics of the training and validation cohort in ovarian cancer

<table>
<thead>
<tr>
<th></th>
<th>Training cohort</th>
<th>Validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of patients</td>
<td>304</td>
<td>380</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>59.5 (30-87)</td>
<td>59 (21-80)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>16 (5.26)</td>
<td>20 (5.26)</td>
</tr>
<tr>
<td>III</td>
<td>242 (79.61)</td>
<td>31 (8.16)</td>
</tr>
<tr>
<td>IV</td>
<td>44 (14.47)</td>
<td>266 (70.00)</td>
</tr>
<tr>
<td>Unknown</td>
<td>2 (0.66)</td>
<td>63 (16.58)</td>
</tr>
<tr>
<td>FIGO stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td>20 (5.26)</td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>31 (8.16)</td>
</tr>
<tr>
<td>III</td>
<td></td>
<td>266 (70.00)</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td>63 (16.58)</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>26 (8.55)</td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td>269 (88.49)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>74 (19.47)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>281 (73.95)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>9 (2.96)</td>
<td>25 (6.58)</td>
</tr>
</tbody>
</table>

**Residual disease**

<table>
<thead>
<tr>
<th>NMD</th>
<th>55 (18.09)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10 mm</td>
<td>141 (46.38)</td>
</tr>
<tr>
<td>11-20 mm</td>
<td>19 (6.25)</td>
</tr>
<tr>
<td>&gt;20 mm</td>
<td>55 (18.09)</td>
</tr>
<tr>
<td>Unknown</td>
<td>34 (11.18)</td>
</tr>
</tbody>
</table>

**Debulking surgery**

<table>
<thead>
<tr>
<th>Optimal</th>
<th>290 (76.32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-optimal</td>
<td>88 (23.16)</td>
</tr>
<tr>
<td>Inoperable</td>
<td>2 (0.52)</td>
</tr>
</tbody>
</table>

**Median survival time (months)**

| 31.5 (1.0-182.7) | 25.7 (0.03-44.2) |

Values are presented as number (range) or number (%).

FIGO, The International Federation of Gynecology and Obstetrics; NMD, no macroscopic disease.
<table>
<thead>
<tr>
<th>Genes</th>
<th>Training cohort</th>
<th>Validation cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>HR (95%CI)</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>Hazard ratio (95%CI)</td>
</tr>
<tr>
<td>IDH2</td>
<td>-0.005</td>
<td>0.995 (0.991-0.999)</td>
</tr>
<tr>
<td>FANCI</td>
<td>-0.089</td>
<td>0.915 (0.857-0.976)</td>
</tr>
<tr>
<td>CXCR4</td>
<td>-0.005</td>
<td>0.995 (0.991-0.998)</td>
</tr>
<tr>
<td>PRAME</td>
<td>-0.007</td>
<td>0.993 (0.986-1.000)</td>
</tr>
<tr>
<td>PJA2</td>
<td>0.042</td>
<td>1.043 (1.014-1.073)</td>
</tr>
<tr>
<td>ECI2</td>
<td>-0.054</td>
<td>0.948 (0.912-0.985)</td>
</tr>
<tr>
<td>AADAC</td>
<td>-0.065</td>
<td>0.937 (0.892-0.984)</td>
</tr>
</tbody>
</table>
HR, hazard ratio; CI, confidence interval; IDH2, isocitrate dehydrogenase; FANCI, FA complementation group I; CXCR4, C-X-C motif chemokine receptor 4; PRAME, PRAME nuclear receptor transcriptional regulator; PJA2, praja ring finger ubiquitin ligase 2; ECI2, enoyl-CoA delta isomerase 2; AADAC, arylacetamide deacetylase.
Supplementary Fig. 1. A workflow of the quality control process of TCGA-OV samples. TCGA-OV, The Cancer Genome Atlas-Ovarian Cancer; GDC, FULL NAME; AWG, FULL NAME; FPKM, fragments per kilobase of transcript per million.

Supplementary Fig. 2. Variance inflation factor (VIF) evaluation plots for the arylacetamide deacetylase (AADAC) signature in the (A) training and (B) validation cohorts. FIGO, The International Federation of Gynecology and Obstetrics.

Supplementary Fig. 3. Kaplan-Meier analysis showed differences in prognosis between the top 25% and the bottom 25% of the arylacetamide deacetylase (AADAC) signature score in ovarian cancer.