Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Integrative Transcriptomic Analysis of a Stemness-Associated Gene Set on Ovarian Cancer Recurrence

Ece Gumusoglu-Acar
Istanbul University
Uygar Yucekaya
Istanbul University
Sude Sultan Ozdemir
Istanbul University
Samet Topuz
Istanbul University
Tuba Gunel
gunel@istanbul.edu.tr

Istanbul University

Research Article

Keywords: ovarian cancer, cancer stem cells, recurrence, SOX2, PROM1

Posted Date: June 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8893995/v1

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

Additional Declarations: No competing interests reported.

Abstract

Ovarian cancer (OC) is a highly lethal gynaecological malignancy, with a five-year survival rate of less than 30%. Although most patients initially respond to platinum-based chemotherapy, disease relapse occurs in approximately 70–80% of cases, largely due to persistence of cancer stem cells (CSCs). In this study, we investigated the expression of a literature-supported set of six stemness-associated genes previously implicated in OC recurrence: ALDH1A1, ALDH1A2, ALDH1A3, NANOG, PROM1 (CD133), and SOX2, in primary OC (POC) (n=421) and recurrent OC (ROC) (n=62). RNA-seq datasets were obtained from “The Cancer Genome Atlas Program (TCGA)”. An ovarian cancer stem cell (OCSC) dataset obtained from the “Gene Expression Omnibus (GEO)” was used as a reference for comparing tumor profiles with the POC and ROC groups. Differential expression between POC and ROC was assessed using statistical tests, and the findings were experimentally validated by qRT-PCR (POC (n = 8); ROC (n = 8)). In the TCGA dataset, SOX2 and PROM1 were upregulated in ROC compared with POC, while ALDH1A1 showed decreased expression. qRT-PCR validation confirmed significant upregulation of SOX2 and PROM1 in recurrent tumors, supporting their association with CSC-related transcriptional programs in recurrent disease. In contrast, ALDH1A1 showed a statistically significant increase in the qRT-PCR cohort, a trend opposite to the TCGA dataset, while ALDH1A3 showed a slight but consistent elevation, and ALDH1A2 remained unchanged. These findings suggest that OC recurrence is associated with expression alterations in transcriptional and metabolic stem cell-related regulators, consistent with the enrichment of OCSC-like features observed in recurrent tumors.

Introduction

OC is a gynecological malignancy with high mortality worldwide, with a five-year survival rate of approximately 30% in advanced-stage cases. This high mortality rate is largely due to its asymptomatic progression in the early stages (International Federation of Gynecology and Obstetrics (FIGO) stages I–II), leading to diagnosis only after the disease has progressed (FIGO III–IV). While initial responses to platinum-based chemotherapy are favorable, up to 80% of patients experience tumor recurrence, which is typically more aggressive and resistant to chemotherapy [1]. This highlights the increasing need to identify the molecular basis of recurrence and CSCs in OC.

A widely accepted hypothesis proposes that CSCs are key contributors of both recurrence and chemotherapy resistance. CSCs are a small tumor cell subpopulation capable of self-renewal, multipotency, and resistance to chemotherapy-induced apoptosis [2]. They have been identified in various cancer types, including OC [3]. In epithelial ovarian carcinoma, OCSCs originate from various anatomical and cellular sources, including resident progenitor-like stem populations in the ovarian surface epithelium and fallopian tube fimbriae, ascites-derived spheroid-forming cells enriched for *CD44⁺/CD24⁻* phenotypes exhibiting enhanced self-renewal, invasion, and chemoresistance, as well as circulating tumor cells and therapy-induced dedifferentiation of non-stem tumor cells, which are associated with tumor relapse, chemotherapeutic failure, and reduced progression-free survival [4,5].

Recent studies indicate that ascites-induced *ALDH⁺CD44⁺* ovarian tumor cell subsets exhibit significantly enhanced CSC-like properties, such as increased sphere formation, self-renewal, metabolic reprogramming, and chemotherapy resistance, compared to primary tumor cells. OSCs share transcriptomic and phenotypic characteristics with embryonic stem cells, implicating them in tumor initiation, relapse, and treatment resistance. Transcription factors are crucial regulators in maintaining the OCSC phenotype by coordinating essential stemness programs and adaptive transcriptional reprogramming in response to therapy. Specifically, core pluripotency factors such as *NANOG, OCT4, SOX2*, and along with EMT-associated regulators like *SNAI1/2* and *ZEB1/2*, regulate self-renewal, metastatic capability, and chemoresistance in ovarian cancer stem cell populations [5].

Transcriptional factors such as *NANOG, PROM1, and SOX2* regulate stem cell-like properties in OC [7]. *NANOG* and *SOX2* maintain pluripotency by repressing differentiation-related programs and reactivating embryonic pathways, and their overexpression has been associated with epithelial-mesenchymal transition and platinum resistance [8,9,10]. *PROM1* marks tumor-initiating cells with self-renewal capacity and is associated with metastasis and recurrence [11]. Members of the *ALDH1A* family regulate stem cell metabolism through aldehyde detoxification and retinoic acid biosynthesis, contributing to redox balance and chemotherapy resistance [4,12]. In OCSC, coordinated activation of these genes supports survival under cytotoxic stress and promotes tumor regrowth after treatment [6].

In this study, we aimed to identify consistent gene expression patterns that could be used as biomarkers or therapeutic targets in relation to recurrence in ovarian cancer. To this end, we performed comparative gene expression analysis of RNA-seq data from primary (POC) and recurrent ovarian cancer (ROC) samples from a GEO-derived ovarian cancer stem cell (OCSC) dataset and TCGA-derived primary (POC) and recurrent ovarian cancer (ROC) samples, focusing on six stem cell-associated genes. In addition to in silico analyses, we assessed the expression profiles of these genes using qRT-PCR in primary and recurrent ovarian cancer patient samples. By integrating large-scale transcriptomic data with experimental validation, this study examines both transcriptional and metabolic stem cell factors to characterize the stem cell-associated features underlying ovarian cancer recurrence and their association with aggressive, drug-resistant tumor phenotypes.

Materials and Methods

In this study, we integrated large-scale RNA-seq datasets from TCGA with an experimentally validated OCSC dataset and qRT-PCR measurements from primary and recurrent OC tissues to evaluate targeted stem cell-associated transcriptional patterns (Fig. 1). This workflow encompasses dataset acquisition, computational preprocessing, differential and correlation analyses, survival modelling, and experimental validation.

Dataset Collection

To investigate the expression patterns of important stem cell-related genes in OC, RNA-seq datasets were obtained from two major public resources: TCGA and GEO. Transcriptomic profiles were quantified

as fragments per kilobase of transcript per million mapped reads (FPKM) from the TCGA cohort, including 421 POC and 62 ROC samples. FPKM was used for consistency with publicly available processed TCGA data. The GEO dataset (accession number GSM4684836) represents a unique RNA-seq profile derived from experimentally validated OCSCs isolated from high-grade serous ovarian cancer (HGSOC) primary tumor cells.

The study focused on a panel of stem cell-associated genes *ALDH1A1, ALDH1A2, ALDH1A3*, *NANOG, PROM1,* and *SOX2,* which are previously reported to be associated with tumor initiation, self-renewal, differentiation blockade, and resistance to chemotherapy in ovarian and other solid malignancies.

Data Preprocessing and Normalization

We preprocessed the RNA-sequencing data to ensure consistency and comparability across POC and ROC groups. The FPKM data for gene expression underwent a structured multi-step workflow before statistical evaluation and biological interpretation.

Genes with particularly low or inconsistent expression (FPKM < 10 in over 80% of samples) were filtered to improve statistical power. This ensured that only transcripts that were biologically expressed in some way were truncated in the known gene panel. After filtering, FPKM values were subjected to $\log_2$ displays [$\log_2$(FPKM+1)] to minimize variance linkages and high expression genes. Differential expression between primary and recurrent tumor groups was assessed using parametric statistical tests on $\log_2$ transformed expression values selected according to data distribution characteristics and unequal sample sizes, followed by false discovery rate (FDR) correction. Genes were considered significantly differentially expressed when |$\log_2$ fold change| > 1 and the corrected p-value < 0.05. The same $\log_2$ transformed expression matrix was used for subsequent analyses, including visualization and exploratory assessments. In addition to this analysis, pairwise correlation analyses were performed among the six target genes to examine shared expression patterns and investigate transcriptional coordination.

Correlation Analysis

Sample-wise correlation analyses were performed to assess the similarity of tumor transcriptomes to the OCSC expression profile. Log_2 transformed expression levels of the genes of interest were used to calculate Pearson correlation coefficients between each tumor sample (primary or recurrent) and the OCSC stem cell signature. This approach quantified how similar each sample was to the OCSC expression profile of the selected genes. Group-level comparisons of correlation coefficients between primary and recurrent tumors were visualized using box plots, and differences were assessed with two-sided t-tests. All analyses and visualizations were performed in Python. Statistical significance was set at $p < 0.05$.

Overall Survival (OS) Analysis

Overall survival was defined as the time from initial diagnosis to death or last follow-up. Clinical and expression data for OC patients were obtained from the TCGA dataset. Cases without OS duration or vital status information were excluded from the analysis.

To assess the prognostic significance of relevant genes, a composite expression score was calculated by averaging the z-score expression values of all six genes for each patient. Kaplan-Meier survival analysis was performed by stratifying patients according to the median of this composite score, and differences in OS were assessed using the log-rank test. Given the exploratory nature of the survival analysis, multivariate relationships between stem cell score and OS were assessed more thoroughly using Cox proportional hazards regression, adjusted for available clinical covariates. Hazard ratios and 95% confidence intervals were reported. All analyses were performed in Python. The statistical significance level was defined as $p < 0.05$.

RNA Isolation and Validation of Gene Expression by qRT-PCR

Clinical samples were collected from patients diagnosed with primary and recurrent epithelial ovarian cancer who were being treated at the Department of Obstetrics and Gynecology Oncology, Istanbul University Faculty of Medicine. POC samples were collected at the time of initial diagnosis, before first-line platinum-based chemotherapy, while ROC samples were collected after documented disease recurrence following standard chemotherapy. All samples were processed under cold chain conditions and stored at -80 °C until analysis. Ethical approval was obtained from the Istanbul University Clinical Research Ethics Committee (Approval No: 2019/242, Date: 22/02/2019) and written informed consent was obtained from all patients.

Total RNA was isolated using the Aurum™ Total RNA Mini Kit (Bio-Rad, USA) according to the manufacturer's protocol. RNA concentration and purity were evaluated using a NanoDrop One spectrophotometer (Thermo Scientific, USA). Complementary DNA (cDNA) was synthesized from a standardized RNA input (100 ng/µL) using the PrimeScript™ RT Reagent Kit (Takara Bio, Japan). Quantitative real-time PCR (qRT-PCR) was performed using SsoAdvanced™ Universal SYBR® Green Supermix on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad, USA). Primer concentrations and reaction conditions were applied identically, and *GAPDH* was used as the internal reference gene. Amplification specificity was confirmed by melting curve analysis.

Results

Differential Gene Expression

Differential expression analysis was performed on a pre-defined panel of six stem cell-related genes to compare POC and ROC samples. Expression differences were assessed using $\log_2$-transformed RNA-seq data from the TCGA cohort. This analysis was designed to evaluate directional and magnitude changes within this targeted stem cell-related gene cluster, rather than a transcriptome-wide scan.

Expression Comparison Across Groups

To compare ROC and POC populations and identify distinctive expression patterns, gene expression analysis was performed, focusing on genes previously associated with CSC biology and treatment resistance. Log_2-transformed expression values were visualized using box plots and bar graphs to assess group differences, variability, and overall trends, and correlation analyses were performed to evaluate common expression patterns among stem cell-associated genes. Box plot analysis revealed that *SOX2* and *PROM1* were significantly upregulated in ROC samples, showing higher median expression levels and wider distributions compared to POC (adjusted $p < 0.01$) (Fig. 2A). These findings were further corroborated by bar graph analysis showing significantly increased *PROM1* and *SOX2* expression in ROC samples (Fig. 2B).

PROM1 was significantly enriched in the recurrent group ($p < 0.001$), whereas *ALDH1A1* expression was significantly decreased in ROC compared to POC ($p < 0.01$). No significant difference was observed for *ALDH1A2, ALDH1A3,* or *NANOG*. Overall, these results demonstrate selective upregulation of stem cell-related key regulators in recurrent tumors.

Correlation Analysis of Stemness-Associated Genes

To investigate transcriptional similarity of stem cell-associated expression patterns among OC cohorts, correlation analyses were conducted between the OCSC, POC, and ROC datasets. The overall stem cell signature obtained from OCSC cells exhibited a negative trend with primary tumors ($r = -0.59$), while its correlation with recurrent tumors was almost neutral ($r = -0.04$) (Fig. 3).

These results indicate that primary ovarian tumors exhibit an opposite expression pattern compared to OCSC, with decreased stem cell-associated activity in early-stage disease. In contrast, recurrent tumors show a partial relative similarity toward an OCSC-like gene expression profile, suggesting that recurrence may be more strongly associated with stem cell programs.

Association of Gene Expression with Overall Survival

Overall survival (OS) analysis was performed using a composite score derived from six stem cell-related genes. Kaplan-Meier analysis showed a trend towards worse overall survival in patients with higher composite expression scores; however, this difference was not statistically significant (log-rank $p = 0.11$) (Fig. 4).

Gene Expression Results in Patient Samples

qPCR was used to measure gene expression levels using the ΔΔCq method in the POC and ROC groups. Relative expression levels were compared among the groups (Fig. 5).

Among the analyzed genes, *ALDH1A1* expression was markedly higher in the recurrent group compared to primary samples ($p < 0.05$). For *ALDH1A2*, expression showed a mild, non-significant increase in the

recurrent group relative to the primary group ($p > 0.05$). *ALDH1A3* levels were significantly higher in the recurrent samples compared to the primary group ($p < 0.05$). Stem cell-associated genes showed variable expression patterns. *SOX2* expression was significantly higher in the recurrent group compared to the POC group ($p < 0.01$). *NANOG* expression showed a moderate but nonsignificant increase in recurrent tumors. Another well-known stem cell marker, *PROM1*, showed a significant increase in the ROC group compared to POC tumors.

A correlation analysis between fold-change values obtained from qPCR and RNA-seq was performed to assess the concordance between the two analytical platforms (Fig. 6). Overall, a weak to moderate positive trend was observed (Spearman $r = 0.42$), indicating partial concordance in gene expression patterns despite biological and technical variability between the datasets.

Discussion

In the present study, integrated analysis of RNA-seq datasets together with qRT-PCR validation demonstrated that ovarian tumors undergo transcriptional reprogramming characterized by activation of stem cell-associated genes. The most reproducible pattern across datasets was the upregulation of *SOX2* and *PROM1* in ROC compared to POC samples in the TCGA cohort. This observation suggests that recurrence and disease progression in OC may be associated with the reactivation of a limited subset of pluripotential regulators, rather than the combined effect of all stem cell markers.

Members of the *ALDH* family exhibited heterogeneous trends, suggesting complexity among subgroups of tumor samples. Although *ALDH1A1/ALDH*-positive cells have been reported to be associated with a stem-like phenotype and chemoresistance [4,13]. In the TCGA RNA-seq dataset, *ALDH1A1* was moderately downregulated in recurrent tumors compared to primary samples. In contrast, in the qRT-PCR cohort, it exhibited a statistically significant increasing trend in recurrent tissues compared to the primary groups. This discrepancy suggests that *ALDH1A1* expression is influenced by chemotherapy exposure, tumor subtype, and microenvironmental factors; this is consistent with the heterogeneous ALDH biology reported in the literature [14,15].

Low *ALDH1A2* levels in OC tumors have been reported to be associated with poor prognosis [16]. In our study, *ALDH1A2* showed a slight but non-significant increase in the recurrent group in both RNA-seq and qRT-PCR analyses, suggesting relatively stable expression during tumor progression. In contrast, *ALDH1A3*, which has been reported to increase tumor aggressiveness [17], exhibited high levels in relapsed tumors in both data sets. The increase reached statistical significance in qRT-PCR analysis ($p < 0.05$) but remained non-significant in the RNA-seq cohort, potentially reflecting sample heterogeneity or limited statistical power. These findings suggest that *SOX2* and *PROM1* transcriptionally significantly define stem cell-like tumor cell programs, while *ALDH* isoforms show heterogeneous regulation rather than uniform activation across tumor states. This suggests that chemotherapy, sample type, and tumor microenvironment, as described in the literature, influence *ALDH* functions. Furthermore, we suggest that stem cell-like programs are primarily mediated by the *SOX2–PROM1* axis.

SOX2 exhibited a consistently strong effect across multiple analyses. RNA-seq data revealed a significant increase in *SOX2* expression in recurrent tumors compared to primary tumors. At the same time, qRT-PCR analysis of patient samples confirmed significant overexpression in the recurrent group compared to primary OC tissues. *SOX2* reactivation is likely associated with EMT, spheroid formation, and treatment tolerance, key processes that facilitate survival after chemotherapy [17,18,19]. These findings are consistent with previous reports identifying *SOX2* as a key factor in EMT, stem cell characteristic, and chemotherapy resistance in ROC. The concordance between our RNA-seq and qRT-PCR data supports *SOX2* as a key regulator of ovarian cancer persistence and recurrence.

Both RNA-seq and qRT-PCR analyses showed a statistically significant increase in *PROM1* expression in recurrent tumors ($p < 0.01$), consistent with the enrichment of the *CD133*$^+$ cancer stem cell-like cell population in the recurrent tumor microenvironment. The concordant elevation of *PROM1* levels in both data sets supports the robustness of this finding. At the transcriptomic level, parallel upregulation of *SOX2* and *PROM1* strongly suggests that a transcriptionally coupled *SOX2*$^+$*/CD133*$^+$ population contributes to tumor persistence and recurrence potential. These findings are in line with previous studies demonstrating that *CD133*$^+$ subsets play a role during ovarian cancer recurrence and are an important population in both chemotherapy resistance and long-term tumor regeneration [6,20].

Correlation analysis further supported these results. Primary tumors exhibited a negative correlation with the OCSC-derived stem cell profile, while recurrent tumors showed that this negative relationship disappeared, approaching a neutral or weakly positive correlation. This shift implies a gradual molecular convergence toward stem cell-like transcriptional states during disease progression. These results suggest that relapse does not originate from a distinct cellular lineage, but rather from tumor cell subsets that gradually regain OCSC-like transcriptional properties [6,21]. Consistent with previous reports, stem cell-associated genes such as *SOX2, NANOG*, and *PROM1* are primarily linked to developmental and self-renewal programs, while members of the *ALDH* family are involved in retinol and aldehyde metabolism. The coexistence of transcriptional stem cell regulators and metabolic enzymes suggests that ovarian cancer recurrence may involve coordinated transcriptional and metabolic adaptation supporting tumor plasticity and chemotherapy resistance [17,22].

Clinically, Kaplan–Meier analysis based on TCGA data showed that patients with higher composite expression of the six-gene stem cell signature tended to have shorter overall survival. However, the difference did not reach statistical significance. Also, this trend suggests that even moderate activation of the combined transcriptional–metabolic program is associated with a more aggressive tumor phenotype. The consistent overexpression of *SOX2* in both RNA-seq and qRT-PCR datasets positions it as a reliable biomarker candidate for relapse-associated stem cell activation [23]. Together, these findings indicate that ovarian cancer progression and relapse are driven by coordinated transcriptional and metabolic reprogramming centred on *SOX2–PROM1* activation and *ALDH* modulation. Thus, the relapse phenotype results from the selective enrichment of preexisting OCSC populations that evade chemotherapy and trigger tumor regrowth. Therefore, targeting this transcriptional stem cell signature

may be an effective therapeutic approach that could eliminate these resistant cell populations and improve long-term treatment outcomes.

Despite the significant results obtained with this integrated approach, some limitations must be considered. The ROC sample size in the TCGA dataset is smaller than that of the primary tumor cohort, which may reduce the ability to detect subtle transcriptional differences. Furthermore, qRT-PCR validation was performed on a limited number of clinical samples, which may somewhat affect the generalizability of the results. Finally, while the six-gene panel provides a focused perspective on stem cell-related alterations, larger multi-omics or functional studies will be needed to fully elucidate the mechanistic links between *SOX2-PROM1* activation, *ALDH* modulation, and chemotherapy resistance in OC.

In conclusion, our findings support a model in which recurrent ovarian tumors are driven by a *SOX2-PROM1* centered stem cell program, accompanied by *ALDH* gene activity that remains heterogeneous across tumor samples. By integrating publicly available RNA-seq datasets with patient-derived qRT-PCR validation, we demonstrated that ROC converges towards a partially OCSC-like transcriptional state, and the activation of this stemness gene set corresponds to a more aggressive phenotype. These observations highlight *SOX2* and *PROM1* as potent markers of recurrence-associated stem cell characteristics and may help guide more precise treatment strategies to prevent chemotherapy-resistant tumor recurrence in OC.

Declarations

Funding: The present study was supported by the Istanbul University Scientific Research Projects Department (grant nos. 42590).

Competing Interests: The authors declare no conflict of interest.

Author Contributions: Conceptualization, Ece Gumusoglu-Acar, Uygar Yucekaya and Tuba Gunel; Formal analysis, Ece Gumusoglu-Acar, Uygar Yucekaya and Tuba Gunel; Investigation, Uygar Yucekaya and Tuba Gunel; Methodology, Uygar Yucekaya and Sude Ozdemir; Project administration, Tuba Gunel; Resources, Samet Topuz; Software, Uygar Yucekaya; Supervision, Tuba Gunel; Validation, Uygar Yucekaya and Sude Ozdemir; Writing – original draft, Ece Gumusoglu-Acar, Uygar Yucekaya, Sude Ozdemir and Tuba Gunel.

Data Availability Statement: The RNA-seq expression and clinical data analyzed in this study were obtained from The Cancer Genome Atlas (TCGA-OV). The ovarian cancer stem cell (OCSC) RNA-seq dataset was retrieved from the Gene Expression Omnibus (GEO) under accession number GSM4684836. All datasets used in this study are publicly available.

Ethics Approval: The study was conducted in accordance with the Declaration of Helsinki and approved by the Istanbul University Clinical Research Ethics Committee (Approval No: 2019/242, Date: 22 February 2019).

Consent to participate: Informed consent was obtained from all subjects involved in the study.

Consent to publish: All authors have reviewed and approved the final version of the manuscript and consent to its publication.

References

1. Havasi A, Cainap SS, Havasi AT, Cainap C. Ovarian cancer insights into platinum resistance and overcoming it. Medicina (Kaunas). 2023;59:544. https://doi.org/10.3390/medicina59030544
2. Aramini B, Masciale V, Grisendi G, Bertolini F, Maur M, Guaitoli G et al. Dissecting tumor growth: the role of cancer stem cells in drug resistance and recurrence. Cancers (Basel). 2022;14:976. https://doi.org/10.3390/cancers14040976
3. Chu X, Tian W, Ning J, Xiao G, Zhou Y, Wang Z et al. Cancer stem cells: advances in knowledge and implications for cancer therapy. Signal Transduct Target Ther. 2024;9:170. https://doi.org/10.1038/s41392-024-01851-y
4. Meng E, Mitra A, Tripathi K, Finan MA, Scalici J, McClellan S et al. ALDH1A1 maintains ovarian cancer stem cell-like properties by altered regulation of cell cycle checkpoint and DNA repair network signaling. PLoS One. 2014;9:e107142. https://doi.org/10.1371/journal.pone.0107142
5. Wilczynski JR, Wilczynski M, Paradowska E. Cancer stem cells in ovarian cancer—a source of tumor success and a challenging target for novel therapies. Int J Mol Sci. 2022;23:2496. https://doi.org/10.3390/ijms23052496
6. Fraszczak K, Zimna A, Nowak D, Czyz J. Ovarian cancer stem cells: molecular features and therapeutic targets. Int J Mol Sci. 2023;24:5523. https://doi.org/10.3390/ijms24065523
7. Robinson M, Gilbert SF, Waters JA, Lujano-Olazaba O, Lara J, Alexander LJ et al. Characterization of SOX2, OCT4 and NANOG in ovarian cancer tumor-initiating cells. Cancers (Basel). 2021;13:262. https://doi.org/10.3390/cancers13020262
8. Roy L, Bobbs A, Sattler R et al. CD133 promotes adhesion to the ovarian cancer metastatic niche. Cancer Growth Metastasis. 2018;11:1179064418767882. https://doi.org/10.1177/1179064418767882
9. Liu S, Sun J, Cai B et al. NANOG regulates epithelial–mesenchymal transition and chemoresistance through activation of the STAT3 pathway in epithelial ovarian cancer. Tumour Biol. 2016;37:9671–9680. https://doi.org/10.1007/s13277-016-4848-x
10. Wen Y, Hou Y, Huang Z, Cai J, Wang Z. SOX2 is required to maintain cancer stem cells in ovarian cancer. Cancer Sci. 2017;108:719–731. https://doi.org/10.1111/cas.13186
11. Kryczek I, Liu S, Roh M, Vatan L, Szeliga W, Wei S et al. Expression of aldehyde dehydrogenase and CD133 defines ovarian cancer stem cells. Int J Cancer. 2012;130:29–39. https://doi.org/10.1002/ijc.25967
12. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. Oncotarget. 2016;7:11018–11032. https://doi.org/10.18632/oncotarget.6920

13. Nwani NG, Condello S, Wang Y, Swetzig WM, Barber E, Hurley T et al. A novel ALDH1A1 inhibitor targets cells with stem cell characteristics in ovarian cancer. Cancers (Basel). 2019;11:502. https://doi.org/10.3390/cancers11040502
14. Landen CN Jr, Goodman B, Katre AA, Steg AD, Nick AM, Stone RL et al. Targeting aldehyde dehydrogenase cancer stem cells in ovarian cancer. Mol Cancer Ther. 2010;9:3186–3199. https://doi.org/10.1158/1535-7163.MCT-10-0563
15. Uddin MH, Kim B, Cho U, Azmi AS, Song YS. Association of ALDH1A1–NEK-2 axis in cisplatin resistance in ovarian cancer cells. Heliyon. 2020;6:e05442. https://doi.org/10.1016/j.heliyon.2020.e05442
16. Choi JA, Kwon H, Cho H, Chung JY, Hewitt SM, Kim JH. ALDH1A2 is a candidate tumor suppressor gene in ovarian cancer. Cancers (Basel). 2019;11:1553. https://doi.org/10.3390/cancers11101553
17. Huang J, Tang Y, Li Y, Wei W, Kang F, Tan S et al. ALDH1A3 contributes to tumorigenesis in high-grade serous ovarian cancer by epigenetic modification. Cell Signal. 2024;116:111044. https://doi.org/10.1016/j.cellsig.2024.11104
18. Mirzaei S, Paskeh MDA, Entezari M, Mirmazloomi SR, Hassanpoor A, Aboutalebi M et al. SOX2 function in cancers: association with growth, invasion, stemness and therapy response. Biomed Pharmacother. 2022;156:113860. https://doi.org/10.1016/j.biopha.2022.113860
19. Seo E, Kim D, Jang I, Choi E, Shin S, Lee S et al. Hypoxia–NOTCH1–SOX2 signaling is important for maintaining cancer stem cells in ovarian cancer. Oncotarget. 2016;7:55624–55638. https://www.oncotarget.com/article/10954/text/
20. Liu CL, Chen YJ, Fan MH, Liao YJ, Mao TL. Characteristics of CD133-sustained chemoresistant cancer stem-like cells in human ovarian carcinoma. Int J Mol Sci. 2020;21:6467. https://doi.org/10.3390/ijms21186467
21. Ding J, Zhang Y, Che Y. Ovarian cancer stem cells: critical roles in anti-tumor immunity. Front Genet. 2022;13:998220. https://doi.org/10.3389/fgene.2022.998220
22. Silva IA, Bai S, McLean K, Yang K, Griffith K, Thomas D et al. Aldehyde dehydrogenase in combination with CD133 defines angiogenic ovarian cancer stem cells that portend poor patient survival. Cancer Res. 2011;71:3991–4001. https://doi.org/10.1158/0008-5472.CAN-10-3175
23. Mamun MA, Mannoor K, Cao J, Qadri F, Song X. SOX2 in cancer stemness: tumor malignancy and therapeutic potentials. J Mol Cell Biol. 2020;12:85–98. https://doi.org/10.1093/jmcb/mjy080

Figures

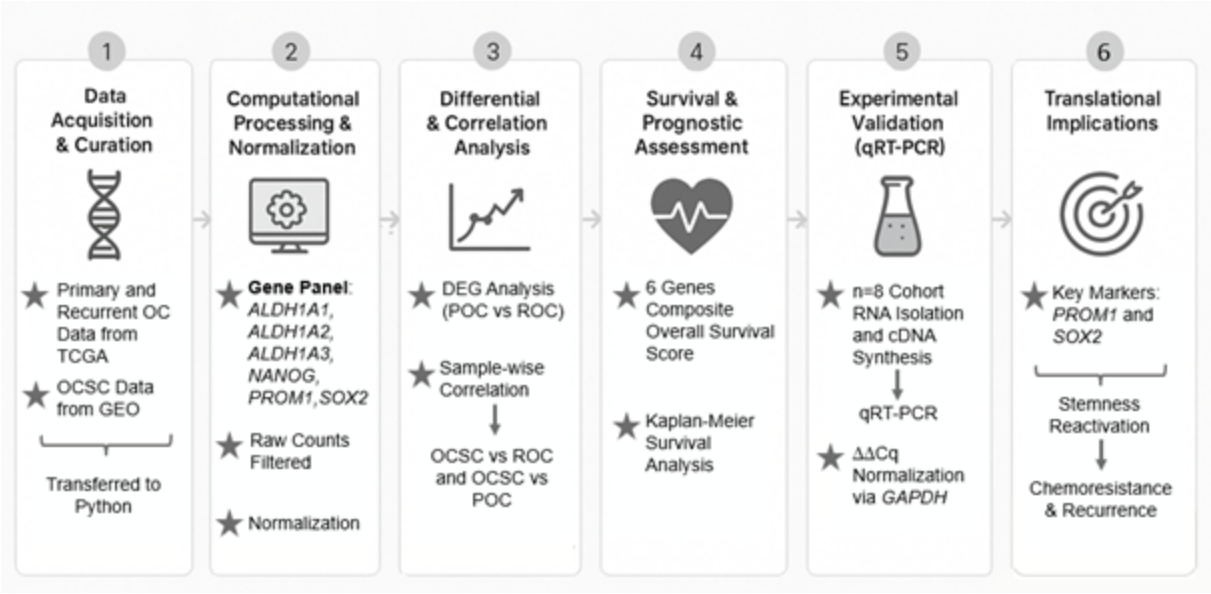


Figure 1

Schematic workflow summarizing the experimental and computational steps of the study

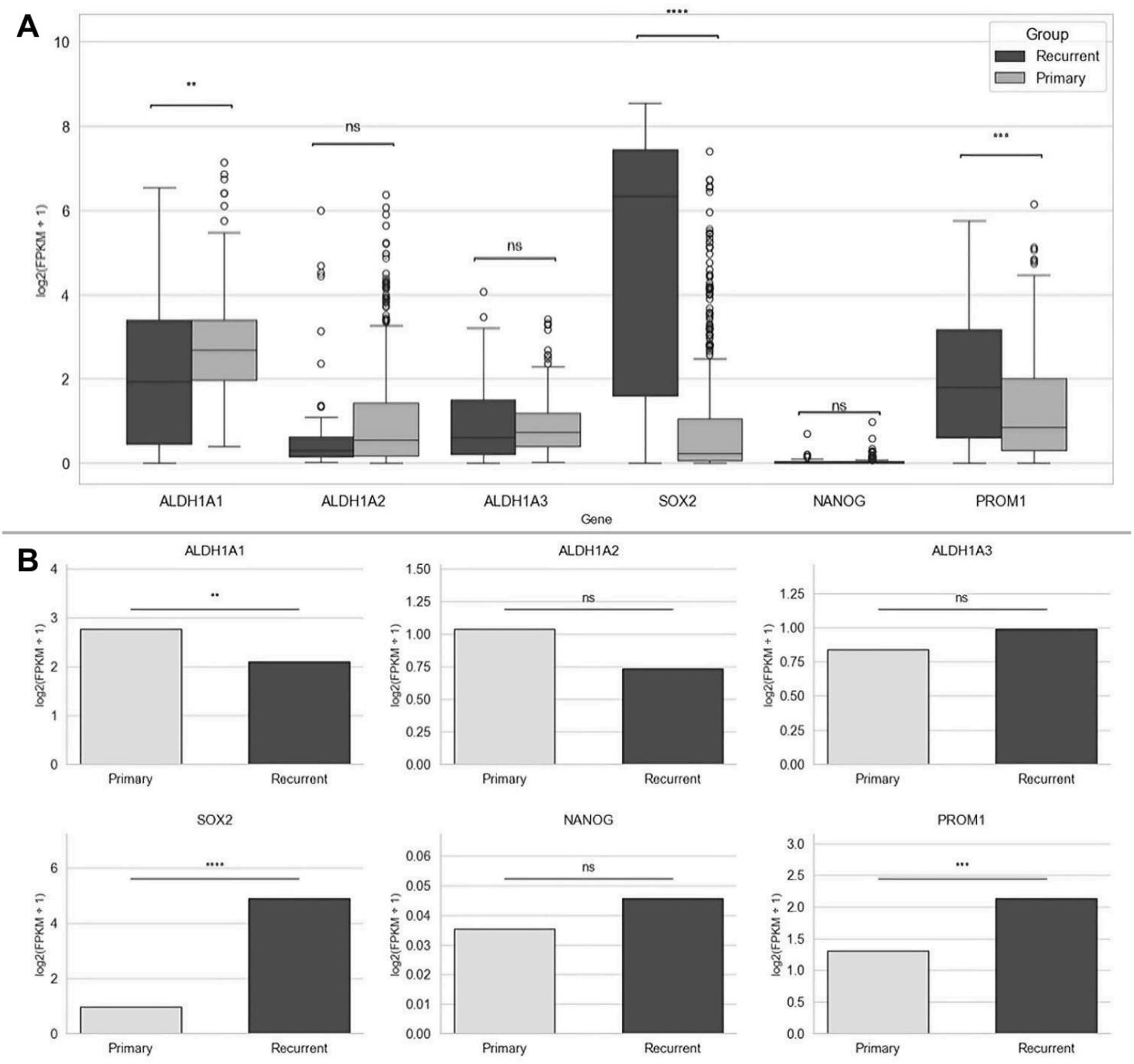


Figure 2

A. Gene expression distribution of stemness-associated genes between POC and ROC samples
B.Differential expression of stemness-related genes between POC and ROC samples. Recurrent tumors showed significant upregulation of *SOX2* and PROM1, while *ALDH1A1* was downregulated. Other genes showed no significant change

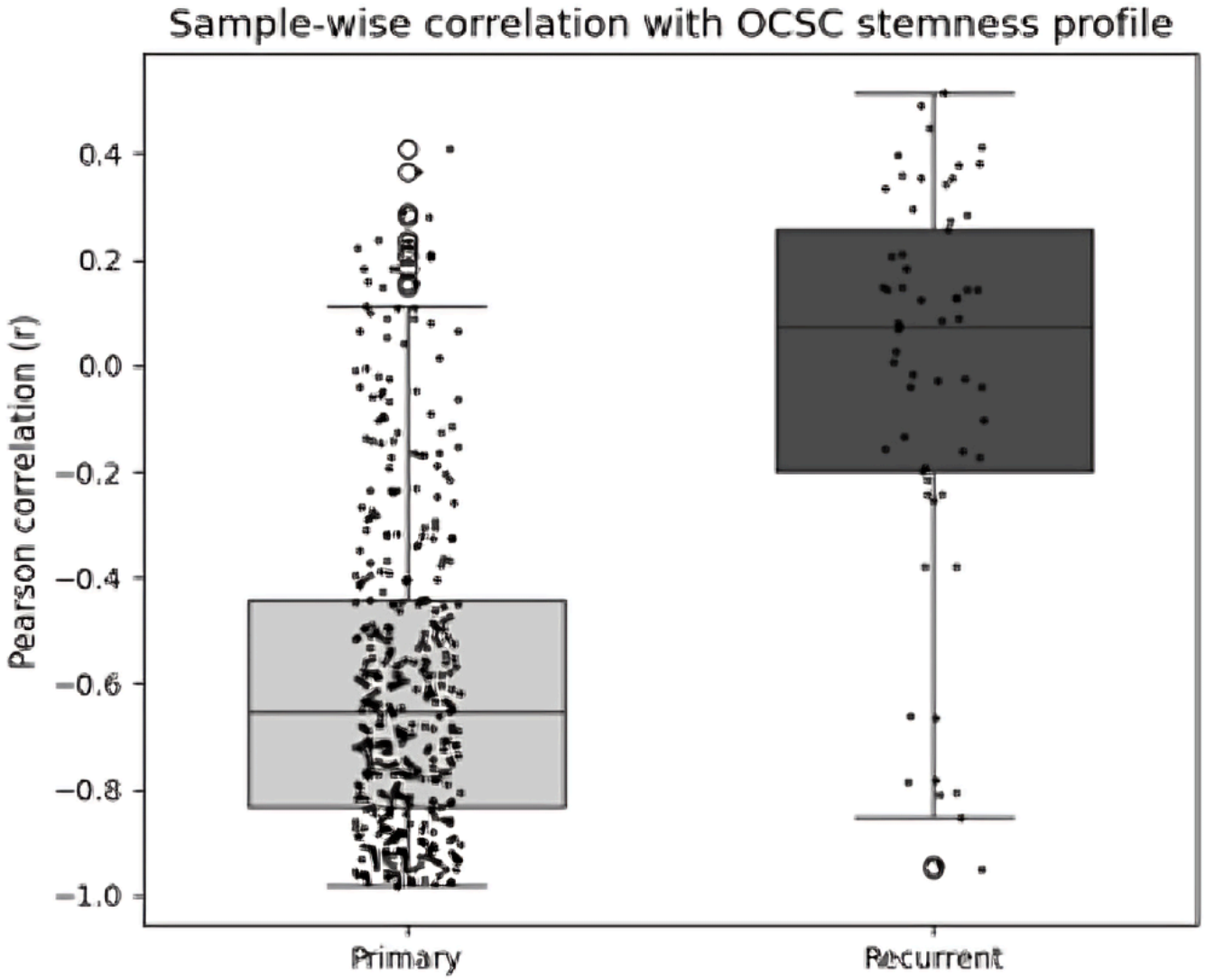


Figure 3

Sample-wise correlation of OCSC stemness profile with primary and recurrent tumors. Recurrent samples showed higher correlation with the OCSC stemness signature, while primary tumors exhibited mostly negative correlations

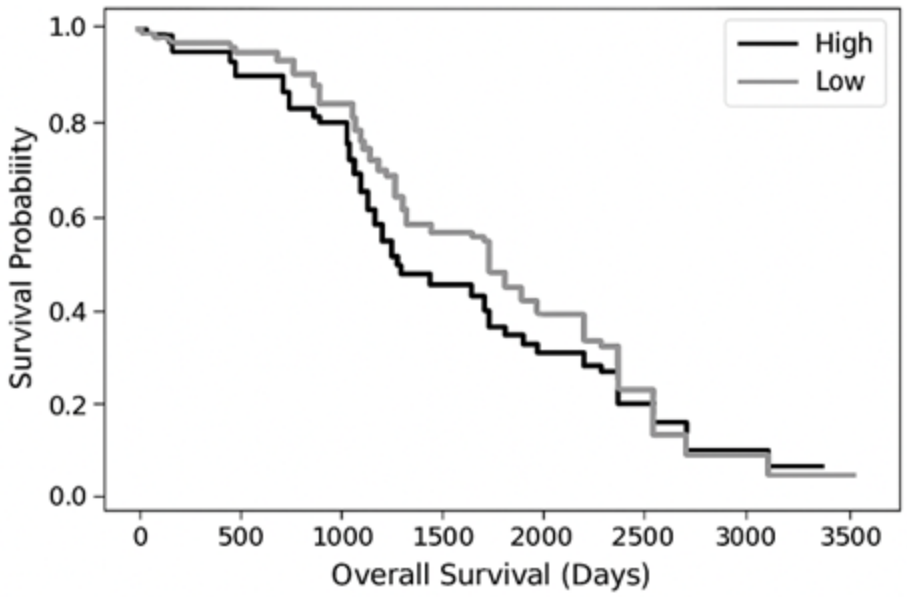


Figure 4

Kaplan–Meier survival curves comparing OS between high and low stemness gene expression groups in OC patients

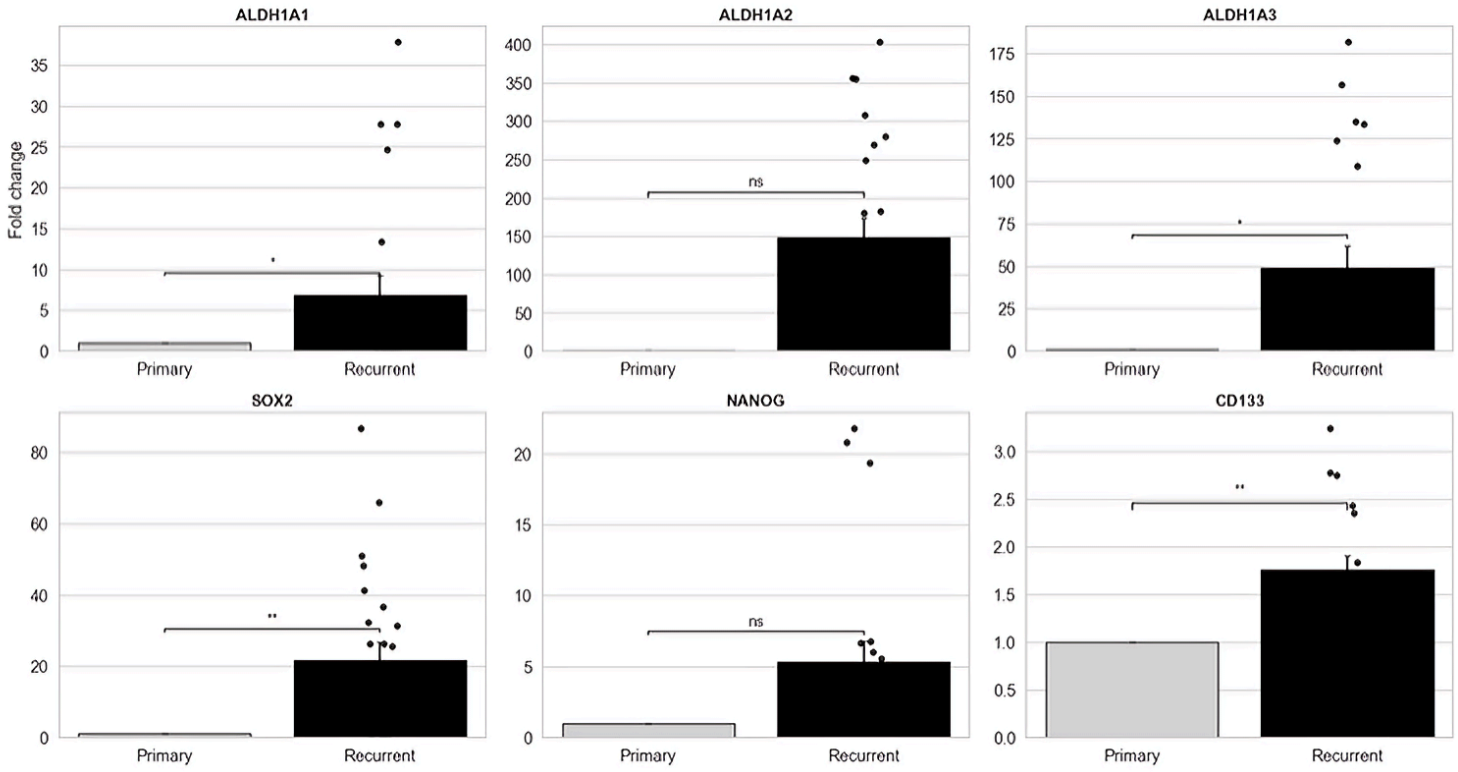


Figure 5

qRT-PCR validation of stemness-associated genes in primary and recurrent OC tissues

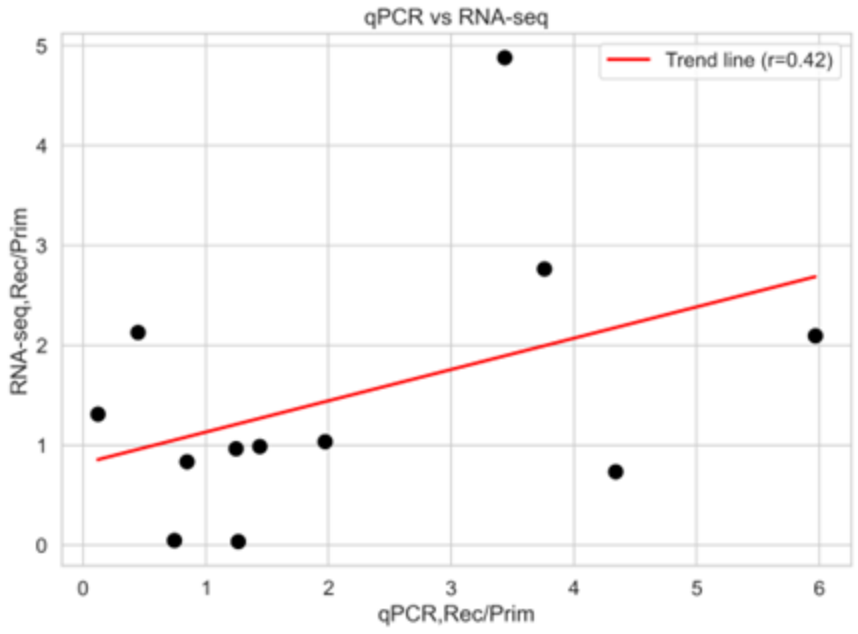


Figure 6

Correlation of the ROC/POC ratio of the genes of interest between qRT-PCR and RNA-seq results