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.

Integrative transcriptomic and network-guided functional dependency analysis reveals essential transcriptional drivers in ovarian cancer

Joel Ruiz-Hernández
Universidad Nacional Autónoma de México
Daniel Pérez-Calixto
National Institute of Genomic Medicine
José Manuel Pastelín-Morales
Universidad del Papaloapan
Enrique Hernández-Lemus
National Institute of Genomic Medicine
Genaro Vazquez-Victorio
Universidad Nacional Autónoma de México
Angélica S. Martínez-Ramírez
Universidad del Papaloapan
Hugo Tovar
hugo.tovar@ciencias.unam.mx
Universidad Nacional Autónoma de México

Research Article

Keywords: Ovarian cancer, Transcriptomic meta-analysis, Gene regulatory networks, Master regulator analysis, Functional dependency, Therapeutic targets

Posted Date: August 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10669238/v1

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

Additional Declarations: No competing interests reported.

Integrative transcriptomic and network-guided functional dependency analysis reveals essential transcriptional drivers in ovarian cancer

Joel Ruiz-Hernández[1,2], Daniel Pérez-Calixto[2,3], José Manuel Pastelín-Morales[4], Enrique Hernández-Lemus[2], Genaro Vazquez-Victorio[3,5], Angélica S. Martínez-Ramírez[6*], Hugo Tovar[3,2*]

[1]Postgraduate Program in Biochemical Sciences, National Autonomous University of Mexico, Mexico City, 04510, Mexico.
[2]Population Genomics Division, National Institute of Genomic Medicine, Mexico City, 14610, Mexico.
[3]Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico.
[4]Division of Postgraduate Studies, MSc Program in Biotechnology, University of Papaloapan, Tuxtepec, 68301, Oaxaca, Mexico.
[5]National Laboratory of Biomimetic Solutions for Diagnosis and Therapy, University City, Mexico City, 04510, Mexico.
[6]Scientific Research Center, Institute of Biotechnology, University of Papaloapan, Tuxtepec, 68301, Oaxaca, Mexico.

*Corresponding author(s). E-mail(s): asmartinez@unpa.edu.mx; hugo.tovar@ciencias.unam.mx;

Abstract

Ovarian serous cystadenocarcinoma remains the most lethal gynecological malignancy, yet its upstream regulators are incompletely defined, hampered by scarce matched normal tissue and limited reproducibility of single-cohort signatures. We integrated transcriptomic evidence from The Cancer Genome Atlas, six Gene Expression Omnibus datasets, and Genotype-Tissue Expression normal ovarian tissue to generate a consensus ovarian cancer signature. Regulatory network analysis identified upstream transcriptional modulators underlying this signature, including transcription factors, cofactors, chromatin remodelers, and epigenetic

regulators. A functional layer integrating CRISPR-Cas9 dependency data from the Cancer Dependency Map distinguished transcriptionally inferred regulators from those genuinely essential for tumor-cell survival. This strategy prioritized 17 central transcriptional drivers that are simultaneously master regulators of essential gene programs and required for ovarian cancer cell viability. These converge on cell-cycle progression, mitotic control, and the DNA-damage checkpoint, including *PLK1*, *CDK1*, *AURKB*, and *CHEK1*, and on chromatin and RNA-processing regulators such as *ACTL6A*, and *SRSF2*; several are actionable or chemo-sensitizing targets, whereas others, including *KCMF1*, *GMNN*, *CHAF1B*, *WDHD1*, and *ELOC*, remain underexplored in ovarian cancer. By coupling regulatory-network inference with functional essentiality, our analysis delivers a prioritized, biologically interpretable set of drivers and therapeutic hypotheses for experimental validation.

Keywords: Ovarian cancer, Transcriptomic meta-analysis, Gene regulatory networks, Master regulator analysis, Functional dependency, Therapeutic targets

Abbreviations

AE, autoencoder-derived control cohort; ARACNe-AP, algorithm for the reconstruction of accurate cellular networks with adaptive partitioning; BH, Benjamini–Hochberg; BP, biological process; CC, cellular component; CRISPR, clustered regularly interspaced short palindromic repeats; DepMap, Cancer Dependency Map; EMT, epithelial-mesenchymal transition; GEO, Gene Expression Omnibus; GO, Gene Ontology; GTEx, Genotype-Tissue Expression; HD-genes, highly dependent genes; HGNC, HUGO Gene Nomenclature Committee; HGSOC, high-grade serous ovarian carcinoma; HRD, homologous recombination deficiency; MF, molecular function; MI, mutual information; MR, master regulator; MRA, master regulator analysis; MSigDB, Molecular Signatures Database; NES, normalized enrichment score; ORA, over-representation analysis; PARP, poly(ADP-ribose) polymerase; RRA, robust rank aggregation; TARGET, Therapeutically Applicable Research to Generate Effective Treatments; TCGA, The Cancer Genome Atlas; TD, transcriptional driver; TF, transcription factor; VIPER, virtual inference of protein activity by enriched regulon analysis.

1 Introduction

Ovarian cancer remains one of the most lethal gynecological malignancies. Although it ranks 11th in frequency among women, it is the 5th leading cause of cancer-related death, largely due to late diagnosis and the absence of specific early symptoms (Bray et al. 2024; Caruso et al. 2025; Dochez et al. 2019). Most patients are diagnosed at advanced stages (III/IV), where the 5-year survival rate drops to 20-40% (Sehgal 2019; Moss et al. 2005; Taş et al. 2019). High-grade serous ovarian carcinoma (HGSOC) accounts for 70-80% of cases and is associated with particularly poor outcomes, while non-epithelial subtypes have a more favorable prognosis (Garlisi et al. 2024). Despite

initial responses to cytoreductive surgery and carboplatin/paclitaxel chemotherapy, 70% of patients eventually relapse with chemoresistant disease (Liu et al. 2025b; Wang et al. 2025). In this context, identifying novel or synergistic therapeutic strategies is essential.

Ovarian cancer is not a single disease but rather a collection of molecularly distinct subtypes, each characterized by different genomic landscapes, transcriptional programs, and clinical behaviors (Amos and Choi 2021; Geistlinger et al. 2020). This heterogeneity manifests as differences in tumor microenvironment composition, immune infiltration patterns, and activation of distinct oncogenic signaling pathways across patients and even within lesions of the same patient (Millstein et al. 2020). As a consequence, molecular markers identified in one patient population frequently lack generalizability, and prognostic signatures derived from a single cohort often fail to stratify risk effectively in independent datasets (Glaab et al. 2021). The absence of robust, widely applicable biomarkers remains one of the central obstacles to improving early detection, guiding treatment selection, and predicting clinical outcomes (Michiels et al. 2016).

The widespread availability of high-throughput transcriptomic data from public repositories such as the Gene Expression Omnibus (GEO) (Edgar et al. 2002) and The Cancer Genome Atlas (TCGA) (The Cancer Genome Atlas Research Network 2011) has enabled genome-wide expression analyses across large patient cohorts (Ganzfried et al. 2013; Tovar et al. 2015; Ruiz-Hernández et al. 2025). These resources have driven efforts to characterize molecular differences between tumor and normal ovarian tissue and to derive gene expression signatures with diagnostic and prognostic value, underscoring the potential of transcriptomics to capture clinically relevant biological variation (Geistlinger et al. 2020; Millstein et al. 2020; Waldron et al. 2014). However, findings from individual cohort studies have shown limited reproducibility across independent datasets, arising from differences in sample composition and tumor grade, heterogeneity in tissue procurement and processing, and the use of distinct technological platforms and normalization strategies (Lazar et al. 2013). As a consequence, differentially expressed genes identified in a single study frequently fail to validate in independent cohorts (Waldron et al. 2014), and meta-analytic approaches have emerged as a principled solution to this problem, aggregating evidence across multiple independent studies to distinguish consistently dysregulated genes from dataset-specific noise (Kolde et al. 2012).

Even when robust meta-analytic signatures are identified, they represent only a first step toward understanding the molecular drivers of disease. Differential expression analysis, while informative, identifies genes whose expression is altered in disease but does not distinguish between upstream drivers and downstream effectors of the dysregulated transcriptional state (Alvarez et al. 2016; Califano and Alvarez 2017). Notably, transcriptional modulators (TM) may exhibit only modest expression changes while exerting broad regulatory influence over gene expression programs (Tovar et al. 2015). Network-based approaches address this limitation by reconstructing regulatory relationships from large-scale data, enabling the identification of central drivers underlying transcriptional dysregulation (Califano and Alvarez 2017). Master Regulator Analysis

(Carro et al. 2010; Lefebvre et al. 2010) builds on these inferred networks to prioritize regulators whose target gene sets are coherently and significantly enriched in a given expression signature, providing a mechanistically interpretable layer of analysis that moves beyond gene lists toward the identification of actionable regulatory drivers of disease. Canonical MRA implementations typically restrict the candidate regulator space to sequence-specific transcription factors. In this work, we extended this framework to encompass a broader class of transcriptional modulators, including transcriptional cofactors, chromatin remodelers, and epigenetic regulators, that, while lacking direct DNA-binding activity, exert pervasive control over gene expression programs through their participation in transcriptional complexes and chromatin state modulation (Lambert et al. 2018). This expanded definition acknowledges that master regulators of disease phenotypes need not be DNA-binding proteins, and that restricting the analysis to canonical transcription factors may systematically overlook key upstream orchestrators of transcriptional dysregulation.

Identifying central transcriptional modulators through network-based approaches yields candidates whose regulatory activity is inferred solely from expression data (Shen et al. 2025a; Lachmann et al. 2016). Orthogonal functional evidence is therefore essential to assess whether these regulators represent genuine dependencies in tumor cells rather than passive correlates of the dysregulated transcriptional state (Tudose et al. 2023). Large-scale CRISPR–Cas9 loss-of-function screens, as compiled in the Cancer Dependency Map (DepMap, Broad 2026) (Tsherniak et al. 2017; Arafeh et al. 2025), provide a systematic and quantitative measure of gene essentiality across hundreds of cancer cell lines, enabling the functional prioritization of candidate regulators based on their contribution to tumor cell survival (Meyers et al. 2017). Integrating network-inferred regulatory activity with CRISPR-derived dependency data thus offers a powerful strategy to identify master regulators that are not only transcriptionally active but also functionally required in ovarian cancer cells.

In the present study, we propose an integrative computational framework that combines multi-source meta-analytic integration, regulatory network inference, and functional dependency evaluation. This approach enables the identification of the master transcription modulators that actively orchestrate the dysregulated transcriptional programs characteristic of ovarian cancer, and assesses their functional essentiality in tumor cell viability. By integrating these complementary layers of evidence, this framework provides a more mechanistically grounded and reproducible strategy for prioritizing molecular drivers and candidate therapeutic targets.

2 Materials and methods

2.1 Overview of the integrative analytical framework

To systematically dissect the transcriptional drivers underlying ovarian cancer, we implemented an integrative strategy that combines cross-cohort transcriptomic meta-analysis, regulatory network inference, and functional dependency assessment (Fig. 1). This framework provides a unified structure to integrate heterogeneous transcriptomic signals and translate them into a coherent regulatory landscape of the disease. Differential expression signals from six GEO microarray datasets and two complementary

RNA-seq contrasts were integrated to derive a high-confidence consensus signature of consistently deregulated genes, capturing robust transcriptional alterations associated with the disease state. This signature was subsequently used to identify key transcriptional drivers within an ovarian tumor-specific transcriptional network, revealing upstream modulators that orchestrate the observed transcriptional program rather than merely reflecting downstream effects. Finally, functional dependency data from DepMap were incorporated to prioritize regulators that are not only transcriptionally active but also essential for ovarian cancer cell viability, thereby refining the candidate set toward biologically and therapeutically relevant drivers.

2.2 Data retrieval and preprocessing

Publicly available ovarian cancer transcriptomic datasets were obtained from the Gene Expression Omnibus (GEO) (Edgar et al. 2002) repository using the *GEOquery* R package (Davis and Meltzer 2007). Six independent studies generated on the Affymetrix Human Genome U133 Plus 2.0 platform (GPL570) were selected (GSE14407, GSE18520, GSE27651, GSE38666, GSE40595, and GSE54388). For each dataset, the corresponding ExpressionSet object was retrieved using the GEO series matrix files, prioritizing GPL570 when multiple platforms were available.

In parallel, publicly available RNA-seq data from ovarian cancer and normal tissues were obtained from three main datasets: The Cancer Genome Atlas (TCGA) (The Cancer Genome Atlas Research Network 2011), the Genotype-Tissue Expression (GTEx) project (The GTEx Consortium 2020), and the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project (Ma et al. 2018). These data were then analyzed using the *OCTAD* (v3.22) R package (Chekalin et al. 2021; Zeng et al. 2021), which provides uniformly processed transcriptomic profiles generated with the Toil pipeline (Vivian et al. 2017). The tumor cohort consisted of 426 TCGA ovarian samples, and was compared against two independent control cohorts of equal size (n = 88 each): (i) normal ovarian tissue samples from GTEx, and (ii) normal-like samples selected using an autoencoder-based approach from a large compendium that included TCGA, GTEx, and TARGET projects (Zeng et al. 2021, 2019).

All transcriptomic datasets used in this study (GEO microarray cohorts, TCGA, GTEx, and TARGET) are publicly available and de-identified; all ethical approvals and informed consents were obtained by the respective originating consortia, and no additional Institutional Review Board approval was required for this study.

2.3 Differential expression analysis

Differential expression analysis for GEO datasets was performed using the *limma* framework (Ritchie et al. 2015). Expression values were extracted and $\log_2$-transformed when required, and samples not classified as case or control were excluded. Linear models were fitted for each gene using a case–control design matrix, followed by empirical Bayes moderation. Reported statistics included log-fold change and Benjamini–Hochberg (BH) adjusted p-values (Benjamini et al. 2001). Probe-level data were annotated to HGNC gene symbols using *g:Profiler* (*gconvert*) (Kolberg et al.

2020). When multiple probes mapped to the same gene, the probe with the highest B-statistic was selected as representative.

For RNA-3 seq data integrated through the Octad R package, differential expression analysis was performed using DESeq2, based on a negative binomial model with library size normalization and dispersion estimation (Zeng et al. 2021; Love et al. 2014). Unwanted sources of variation were modeled using factor analysis as implemented in RUVSeq (45). Statistical significance was assessed at the gene level and corrected for multiple testing using the BH procedure. Gene identifiers were harmonized to HGNC gene symbols, and duplicate entries were resolved by retaining the observation with the largest absolute Wald statistic.

2.4 Robust rank aggregation meta-analysis

To identify genes consistently deregulated across independent GEO studies, Robust Rank Aggregation (RRA) (Kolde et al. 2012; Võsa et al. 2014) was applied using the RobustRankAggreg R package. Separate analyses were performed for upregulated and downregulated genes, using the full gene universe to estimate expected rankings under a null model of random ordering. p-values were adjusted using the BH procedure, and genes with adjusted p-values ¡ 0.05 were considered significant.

RRA yields gene-level p-values derived from aggregated rankings. Up- and downregulated results were merged and collapsed into a single entry per gene, retaining the direction associated with the lowest adjusted p-value. A signed continuous score was then computed by transforming adjusted p-values into Z-scores using the inverse normal function, with directionality defined by the RRA sign (up = +1; down = –1):

$$Z_{\mathrm{RRA}} = \mathrm{sign}_{\mathrm{dir}} \cdot \Phi^{-1}(1 - p_{\mathrm{adj}}/2) \tag{1}$$

. This transformation yields a continuous effect-like metric that preserves both the strength and direction of deregulation, facilitating downstream integrative analyses.

2.5 Cross-cohort concordance analysis

Consistency between GEO-derived RRA signals and TCGA-based differential expression contrasts (TCGA vs GTEx and TCGA vs AE) was assessed using Spearman correlation under two complementary settings: (i) genome-wide concordance, computed as the correlation between gene ranks across all shared genes; and (ii) high-confidence concordance, evaluated on the top 2,000 genes ranked by RRA significance, correlating signed RRA scores with signed DESeq2 statistics. This analysis was performed to evaluate consistency across independent transcriptomic cohorts and to support the integration of GEO and TCGA-derived signals in subsequent analyses.

2.6 Integration of differential expression evidence across sources

Gene-level differential expression results from GEO-derived RRA and TCGA-based contrasts were integrated. Significantly deregulated genes were defined using consistent statistical and effect size thresholds, and directionality was determined by the sign

of the corresponding statistic. Integration was restricted to genes showing consistent directionality across all three sources. Only genes with concordant regulation in GEO (RRA), TCGA–GTEx, and TCGA–AE contrasts were retained, defining a consensus set of reproducibly deregulated genes across independent cohorts.

2.7 Gene expression variability across overlapping gene sets

Expression variability was quantified using the Fano factor (variance-to-mean ratio) (Schofield and Hahn 2024). computed on normalized RNA-seq counts for each gene in the TCGA–GTEx and TCGA–AE datasets. Gene sets were defined by integrating differential expression results across datasets, yielding subsets of genes grouped by shared membership. GEO data were used solely to define gene set membership and were not included in Fano factor calculations. Differences in dispersion were summarized by comparing the median of the Fano factor distribution across conditions.

2.8 Functional enrichment analysis

Functional enrichment was performed using over-representation analysis (ORA) with *clusterProfiler* (v4.16.0) R package (Yu et al. 2012) on the consensus gene set, using the union of all genes across sources as the universe set. Gene Ontology (GO) enrichment was evaluated across Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) categories. Significance was assessed using the BH procedure (adjusted p-value ¡ 0.05). Redundant terms were reduced using semantic similarity (Wang method; cutoff = 0.3) (Wang et al. 2007). Regulons were subjected to the same ORA approach using the Hallmark gene set collection (Liberzon et al. 2015), retrieved through the msigdbr R package, v26.1.0 (Dolgalev 2026).

2.9 Regulatory network and regulon construction

We performed regulatory inference using a previously generated ARACNe-AP (Lachmann et al. 2016) transcriptional network inferred from TCGA ovarian tumor expression data (300 bootstraps; MI p-value threshold 1×10^{-8}). The network was converted into VIPER (Lefebvre et al. 2010) compatible regulons using the same TCGA tumor expression matrix employed during network inference to ensure consistent feature space and gene identifiers.

2.10 Master regulator analysis

Master Regulator Analysis (MRA) (Carro et al. 2010; Lefebvre et al. 2010) was performed using msVIPER (Alvarez et al. 2016) implementation with the core consensus signature as input. To extend the regulatory scope of the analysis beyond sequence-specific transcription factors, the candidate modulator set was defined using TFcheckpoint 2 (Acencio et al. 2024), a manually curated registry that classifies human transcriptional regulators into transcription factors (TF) and transcriptional cofactors (TFC), the latter encompassing chromatin remodeling complexes, histone-modifying enzymes, and coactivator/corepressor assemblies that lack direct sequence-specific DNA-binding activity yet exert pervasive control over gene expression programs. The

union of both categories (TF ∪ TFC; n = 2,702 unique genes) was used to annotate regulators within the ARACNe-AP–inferred network; microRNA entries present in the TFcheckpoint 2 export were excluded, as these molecules are not represented in the transcriptomic data matrices used for network inference. Hereafter, the full set of evaluated regulators, comprising both canonical transcription factors and transcriptional cofactors, is referred to as transcriptional modulators. Because the core signature is derived from integrated multi-source evidence rather than a single sample-level case–control comparison, statistical significance was estimated using a signature-label permutation null model, preserving the distribution and size of the signature while disrupting regulator–target associations. For each run, msVIPER was computed with a minimum transcriptional neighborhood size of 20 targets and 10,000 permutations using a fixed random seed and parallel computation. Empirical two-sided p-values were computed for each regulator as:

$$p_{\text{emp}} = \frac{1 + \#\{\, |NES_{\text{null}}| \geq |NES_{\text{obs}}| \,\}}{N_{\text{perm}} + 1} \tag{2}$$

where NES_{obs} is the observed normalized enrichment score and NES_{null} the null values across N_{perm} permutations. Empirical p-values were adjusted by BH procedure; regulators were ranked by $|NES|$ and adjusted p-value (p-adj ¡ 0.05, $|NES| \geq 2$). To define a high-confidence transcriptional modulators a data-driven cutoff was applied using an elbow (knee-point) (Guven 2023) criterion on the ranked $|NES|$ distribution of MR, retaining those above the point of maximum curvature.

2.11 Functional gene dependency evaluation in ovarian cancer cell lines

Gene dependency of the core signature was evaluated using genome-wide CRISPR–Cas9 dependency scores from the Cancer Dependency Map (DepMap) (Tsherniak et al. 2017), restricted to ovarian cancer cell lines. Genes were ranked by median dependency score, and a threshold of ≤ -0.5 was applied to define highly dependent genes. This subset was used for over-representation analysis with master regulators to identify regulators significantly associated with highly dependent genes (p ¡ 0.05). These significant transcriptional drivers were then mapped to DepMap dependency scores to evaluate their own essentiality profiles across the analyzed cell lines.

2.12 Literature validation of master regulators in ovarian cancer

To assess whether the prioritized transcriptional regulators had been experimentally or clinically investigated in ovarian cancer, we performed a targeted PubMed literature search. To ensure transparency and reproducibility, we used Boolean operators combining each gene's official HGNC symbol with ovarian cancer–related terms. A representative query was: ovarian cancer OR ovary neoplasm OR ovarian carcinoma OR high-grade serous ovarian carcinoma (AND) ("¡ gene symbol¿"). The search covered studies published in the last 10 years, up to April 2026.

3 Results

3.1 A robust gene expression signature defines core transcriptional alterations in ovarian cancer

The identification of appropriate control samples remains a major challenge in RNA-seq studies of ovarian cancer due to the limited availability of matched normal tissues (Zeng et al. 2019; Collado-Torres et al. 2017). To address this limitation, we developed an integrative three-layer framework that connects complementary sources of transcriptomic data across independent cohorts, in this context, we analyzed uniformly processed RNA-seq samples from the TCGA ovarian serous cystadenocarcinoma (TCGA-OV) cohort (The Cancer Genome Atlas Research Network 2011) together with two complementary control cohorts: normal ovarian tissue samples, obtained from GTEx (The GTEx Consortium 2020), and an additional set of reference controls was defined using an autoencoder-based approach (AE) (Zeng et al. 2021, 2019), which identifies transcriptomically matched healthy tissues across multiple repositories, optimizing comparability with the tumor cohort (Table SI). All RNA-seq datasets (TCGA-OV, GTEx, and AE) were derived from the Toil pipeline, ensuring a consistent analytical framework and reducing technical variability across sources (Vivian et al. 2017). In parallel, six independent GEO microarray datasets were incorporated to provide additional transcriptomic evidence across studies.

Under this framework, differential expression was computed independently for each contrast (TCGA–OV relative to GTEx and AE) and for each GEO dataset, followed by meta-analysis of the GEO studies (see Materials and Methods 3.5). Cross-platform concordance was assessed by correlating the signed RRA meta-analytic signal with DESeq2 statistics from each TCGA contrast. Genome-wide correlations were modest (AE: $\rho = 0.211$; GTEx: $\rho = 0.237$), but restricting to genes with the strongest RRA evidence markedly improved concordance (AE: $\rho = 0.546$; GTEx: $\rho = 0.590$), indicating that robust meta-analytic signals are systematically reproducible across independent cohorts (Fig. S1). To confirm that this improvement was not an artifact of subset size, we generated a null distribution by computing Spearman correlations in 1,000 random subsets of equal size drawn from the full gene set; the observed correlations exceeded all random draws in both contrasts (empirical p ¡ 0.001), demonstrating that the enrichment of cross-platform concordance among high-RRA genes reflects a genuine property of the meta-analytic signal (Fig. S2). This finding supports the integration of differential expression evidence by retaining only genes with consistent directionality across all three sources, without combining expression magnitudes across platforms. The intersection of these gene sets defined a consensus set of 646 consistently differentially expressed genes across datasets (Fig. 2A).

To further assess the robustness of this set, we quantified gene-level expression variability using the Fano factor computed from RNA-seq data (TCGA-OV vs GTEx and AE controls). Comparisons across gene subsets defined by different levels of overlap revealed that the consensus set, corresponding to the intersection of all three independent sources, consistently exhibited lower variability than other subsets (Fig. 2B; Fig. S3), indicating more stable expression patterns and reduced noise.

These genes capture the magnitude and direction of transcriptional changes in tumors, highlighting representative upregulated (e.g., *TPX2*, *UBE2C*, *PRAME*, *KLK7*, *CENPA*, *SLC34A2*, *CLDN3*, *FOLR1*, *SCGB2A1*) and downregulated genes (e.g., *SYNE1*, *GPRASP1*, *ZNF540*, *TRPM3*, *CGNL1*, *CACNB2*, *LAMA2*, *MEG3*) (Fig. 2C). Notably, they exhibit consistent expression patterns that clearly distinguish tumors from control samples (Fig. 2D). In turn, functional characterization revealed enrichment of processes related to chromosome segregation, microtubule binding, DNA-dependent ATP activity, and chromosomal organization (Fig. 2E), indicating that this transcriptional profile reflects coherent biological programs (The Cancer Genome Atlas Research Network 2011).

3.2 Master regulator analysis identifies key transcriptional drivers associated with the core gene signature

Having established a robust consensus signature of consistently dysregulated genes across multiple independent cohorts, we next asked which upstream drivers orchestrate this transcriptional program. To address this, we built an ovarian cancer-specific transcriptional modulator network using ARACNe-AP applied to TCGA tumor expression data, and then converted it into VIPER-compatible regulons. Master Regulator Analysis (MRA) with msVIPER, taking the consensus signature as input, identified 328 high-confidence driven regulators based on a data-driven elbow criterion applied to the ranked distribution of absolute normalized enrichment scores ($|NES|$). Among these, 245 were predicted to be activated and 83 repressed in ovarian tumors relative to controls (Fig. 3).

Among the most strongly activated transcriptional drivers, we identified key cell cycle regulators (*CDK1*, *CCNA2*, *CHEK1*), the mitotic kinase gene *AURKB*, centromere-associated genes (*CENPA*, *CENPU*, *TRIP13*), and the transcription factor gene *E2F1* (Yang et al. 2016; Alexandrova et al. 2020; Pérez-Fidalgo et al. 2020). In contrast, the most strongly repressed regulators included transcriptional repressor (*ZEB1*, *HIC1*, *PRDM6*) and developmental signaling factor genes (*ERG*, *SMAD9*, *DACT1*) (Wang et al. 2024; Singh et al. 2020; Vaicekauskaitė et al. 2025). Collectively, these driver regulators modulate the expression of the consensus signature genes, as their resective target genes are significantly enriched within this set.

3.3 Functional dependency profiling in ovarian cancer cell lines reveals transcriptional drivers associated with essential genes

Transcriptomic landscape derived from patient tumors provides a robust framework for defining gene programs associated with the disease state (Shen et al. 2025a). Nevertheless, not all phenotypic changes are therapeutically accessible (Dang et al. 2017). In this sense, cell line models have been instrumental for understanding tumor biology and identifying potential therapeutic vulnerabilities (Ruiz-Hernández et al. 2025; Mirabelli et al. 2019). To further prioritize elements of the ovarian transcriptional regulatory program with greater therapeutic relevance, we evaluated the functional

dependency of these genes across ovarian cancer cell lines using CRISPR–Cas9 loss-of-function screening data from the Cancer Dependency Map (DepMap) (Tsherniak et al. 2017; Hart et al. 2015).

In this context, we first characterized genes within the core signature exhibiting high dependency scores. Using a threshold of ≤ -0.5, as suggested by standard DepMap criteria (Shi et al. 2024; Jaladanki et al. 2021), we defined a subset of 101 highly dependent genes (HD-genes) (Fig. 4A). Notably, several top-ranked genes in this subset are involved in core cellular processes essential for proliferation and viability. These include regulators of DNA replication and genome integrity (*RRM2*, *CDC45*, *CHEK1*) (Shen et al. 2019; Gui et al. 2021), mitotic progression and cell cycle control (*CDK1*, *AURKB*, *ESPL1*) (Xu et al. 2025), chromatin dynamics and DNA topology (*TOP2A*) (Shen et al. 2019), RNA processing and splicing (*SNRPB*, *PUF60*) (Otto and Sicinski 2017; Li et al. 2023; Zhang et al. 2024a).

To identify the regulatory architecture underlying these HD-genes, we performed over-representation analysis using the regulons previously reported in MRA. This analysis identified 46 transcriptional drivers significantly associated with the HD-gene set (Fig. 4B). Among the top-ranked regulators were *CENPA*, *DEPDC1*, *CDK1*, *CCNA2*, *PLK1*, *BLM*, *EZH2*, *CHEK1*, *CENPU*, each linked to a substantial fraction of highly dependent genes, suggesting that these factors may play a central role in coordinating transcriptional programs associated with essential ovarian cellular functions.

We next examined the dependency profiles of these transcriptional drivers to identify those that, beyond controlling highly dependent genes, are themselves functionally required in ovarian cancer cell lines (Fig. 4C). Applying the same dependency threshold (≤ -0.5), we identified 17 central transcriptional drivers, which we then contextualized through a systematic literature search (Table SII). Many of these drivers have been extensively studied in high-grade serous ovarian carcinoma, and their functions map closely onto the core biology of the disease (Fig. 5). Several act at the G2/M transition to govern mitotic entry and cell-cycle progression: *PLK1* (Shen et al. 2025b; Affatato et al. 2020), *CDK1* (Yang et al. 2016; Xu et al. 2025), *AURKB* (Gao et al. 2025b; Auguste et al. 2020), *CCNA2* (Chen et al. 2025; Tan et al. 2025b) and *CDK2* (Lashen et al. 2024; Sine et al. 2025), whereas *CHEK1* (Parmar et al. 2019; Wen et al. 2021), and *TIMELESS* (Zhang et al. 2024b; Xing et al. 2021), coordinate the DNA-damage checkpoint that preserves replication fidelity under genotoxic stress. At the chromatin level, *ACTL6A* (Zhang et al. 2019; Chen et al. 2023) sustains dynamic chromatin remodeling, a dependency that extends beyond transcription-factor activity and is consistent with the expanded, TFC-inclusive regulator. Centromere identity and kinetochore assembly are represented by *CENPA* (Han et al. 2021; Zhang et al. 2022), whose elevated expression and strong dependency reflect the aneuploidy-prone nature of HGSOC, while survival and apoptotic evasion are anchored by *BIRC5* (Liu et al. 2025a; Wang et al. 2018a), and *HMGB3* (Ma et al. 2023, 2022), both linked to platinum resistance and metastatic progression. Post-transcriptional control is represented by the splicing factor *SRSF2* (Wang et al. 2026, 2018b), which modulates platinum sensitivity through alternative splicing.

In contrast, a subset of these drivers:*CHAF1B* (Sykaras et al. 2021), *ELOC* (Moskovich et al. 2025), *WDHD1* (Zhang and Zhu 2023), *GMNN* (Willis et al. 2016;

Weir et al. 2023), *KCMF1* (Xu et al. 2026; Chen et al. 2026), remains comparatively uncharacterized in ovarian cancer, having been described mainly in pan-cancer or non-ovarian contexts. Their emergence here as both transcriptional drivers and functional dependencies, within a disease-specific analysis, defines a set of novel hypotheses that warrant mechanistic investigation. Together, these findings outline a regulatory framework in which central transcriptional drivers coordinate essential gene programs, providing a prioritized resource for experimental validation in ovarian cancer.

4 Discussion

The transcriptional drivers prioritized here converge predominantly on cell-cycle progression, mitotic control, and the DNA-damage checkpoint, consistent with the proliferative and genomically unstable nature of high-grade serous ovarian carcinoma (Verhaak et al. 2013), and with the convergence of their ARACNe-AP regulons on E2F-target and G2/M-checkpoint transcriptional programs. The mitotic kinase *PLK1*, the cyclin-dependent kinases *CDK1* and *CDK2*, cyclin A2 (*CCNA2*), and the aurora kinase *AURKB* coordinate the G1/S and G2/M transitions and are recurrently overexpressed in ovarian carcinoma, where their activity tracks with high grade, chemoresistance, and poor survival (Xu et al. 2025; Shen et al. 2025b). Notably, several of these nodes are already being pursued therapeutically: *PLK1* inhibition with onvansertib resensitizes platinum-resistant tumors (Affatato et al. 2020), *CDK1* inhibition with RO-3306 suppresses tumor growth in cell and mouse models (Huang et al. 2023), and *CHEK1*, overexpressed after chemotherapy to repair genotoxic damage, is targeted by prexasertib (Nair et al. 2020); *CDK2* further couples cell-cycle deregulation to homologous-recombination dependency in *CCNE1*-amplified tumors (Brown et al. 2023). *TIMELESS*, a replication-fork guardian overexpressed in advanced-stage disease and associated with poor overall survival, complements this checkpoint axis as a downstream target of *MEX3A* (Xing et al. 2021; Li et al. 2022); similarly, the centromeric histone *CENPA*, required for kinetochore assembly and linked to poor prognosis and *MYBL2*-dependent regulation, reflects the aneuploidy-prone character of this disease (Dou et al. 2022). That this druggable mitotic axis was recovered through an orthogonal network-plus-dependency strategy reinforces both the biological validity of the framework and the standing of these regulators as actionable vulnerabilities.

The prioritized set also includes regulators acting at the level of chromatin organization, RNA processing, and cell survival, directly validating the extension of the candidate space beyond sequence-specific transcription factors. *ACTL6A*, an oncogenic SWI/SNF subunit, promotes glycolysis, proliferation, and epithelial-mesenchymal transition (Chen et al. 2023); the CAF-1 histone chaperone subunit *CHAF1B* marks highly proliferative tumors and supports DNA repair (Sykaras et al. 2021); and the splicing factor *SRSF2* promotes platinum resistance through modulation of caspase-8/9 splicing (Pellarin et al. 2020). *BIRC5* (survivin), a potent apoptosis inhibitor whose CRISPR-mediated depletion impairs proliferation and induces apoptosis (Zhao et al. 2022), and the chromatin-associated protein *HMGB3*, which drives migration, invasion, and EMT and is linked to lymph-node metastasis (Ma et al. 2023, 2022),

further anchor the survival and plasticity programs of high-grade serous disease. Of particular interest are drivers that remain underexplored in ovarian cancer yet emerged as both transcriptionally active and functionally essential: *KCMF1*, an E3-ubiquitin-ligase activating β-catenin signaling whose silencing reduces tumor growth (Chen et al. 2026); the replication-licensing factor geminin (*GMNN*), overexpressed in high-grade serous tumors (Weir et al. 2023); *WDHD1*, which sustains survival under platinum exposure (Zhang and Zhu 2023); and *ELOC*, implicated in protein-degradation stress responses (Moskovich et al. 2025). Their co-emergence in a disease-specific analysis nominates them as novel, testable hypotheses for mechanistic and preclinical investigation.

Beyond their individual relevance, these results point to combinatorial vulnerabilities that align with current therapeutic strategies in ovarian cancer. *PLK1* and *CHEK1*, both prioritized here, sensitize BRCA-deficient and even homologous recombination-proficient tumors to PARP inhibitors, suggesting that these drivers may help extend PARP-inhibitor benefit beyond the canonical HRD population (Gasimli et al. 2022; Smith et al. 2024). Similarly, pharmacological inhibition of survivin (*BIRC5*) with YM155 resensitizes ovarian cancer cells to paclitaxel, supporting a role for this driver in overcoming taxane resistance (Fan et al. 2024). The chromatin and splicing regulators identified here (*ACTL6A*, *SRSF2*) are not yet targeted by approved agents in ovarian cancer, but their mechanistic link to chemoresistance and proliferation positions them as candidates for epigenetic or splicing-modulator combination strategies (Vaicekauskaitė et al. 2025; Sykaras et al. 2021; Pellarin et al. 2020; Xiao et al. 2021). Likewise, the underexplored drivers *KCMF1*, *CHAF1B*, *GMNN*, *WDHD1*, and *ELOC* currently lack dedicated clinically established inhibitors, and their therapeutic tractability remains to be established (Chen et al. 2025; Xu et al. 2026; Zhao et al. 2022; Lliberos et al. 2024). Overall, the convergence of transcriptional activity and functional essentiality identified by this framework offers a rational basis for prioritizing combination regimens and repurposing candidates for experimental testing in ovarian cancer models.

Differential expression analysis successfully identifies genes that change in tumors, but it does not reveal which of them act as upstream orchestrators of the transcriptional program (Kerslake et al. 2023). MRA complements this by adding a higher regulatory layer through co-expression inference (Tovar et al. 2015; Ferreira et al. 2018). Extending the candidate space to transcriptional modulators, including cofactors, chromatin remodelers, and epigenetic regulators, is particularly consequential, as key upstream drivers are not necessarily DNA-binding proteins and would be missed by a TF-restricted analysis (Acencio et al. 2024; Tan et al. 2025a); an additional advantage is that MRA does not rely on the regulator's own expression level, avoiding the blind spot of drivers with modest fold changes. Nevertheless, inferred interactome neighborhoods represent statistical co-expression dependencies and require experimental validation before causality can be established (Micoli et al. 2025).The integration of DepMap data adds an orthogonal functional layer: requiring regulators to be both transcriptionally active and essential in cell lines provides a more stringent filter than dependency scores alone, avoiding general housekeeping genes (Shi et al. 2024; Creighton 2023). Yet DepMap profiles come from immortalized lines lacking the tumor

microenvironment; essentiality in vitro may not fully translate in vivo, and some dependencies may only emerge in specific contexts (Behan et al. 2019; Gao et al. 2025a). The prioritized list should therefore be viewed as hypothesis-generating rather than conclusive.

A single-cohort comparison between tumor and normal tissue inevitably suffers from multiple sources of bias: GTEx samples alone introduce postmortem degradation artifacts (Padi and Quackenbush 2015), whereas autoencoder-selected controls, despite their transcriptomic matching, aggregate diverse tissues that preclude perfect biological equivalence (Zeng et al. 2019). The integrative framework reported here was designed to overcome these individual shortcomings: by combining meta-analysis of multiple GEO datasets with two independent control strategies and requiring consistent directionality across all three sources, uniform processing with the Toil pipeline (Vivian et al. 2017) and hidden-variation modeling with RUVSeq (Risso et al. 2014), we obtain a consensus signature that is robust, reproducible, and unlikely to be driven by any single dataset or control type. Some limitations nonetheless apply. As with any bulk RNA-seq approach, our analysis averages expression across cell populations and cannot resolve intratumoral heterogeneity (Jeong et al. 2025) stratification by clinical subtype or tumor stage was precluded by insufficient sample sizes within individual strata, with available samples skewed toward advanced disease (Sereti et al. 2025). Nevertheless, bulk transcriptomic approaches have consistently captured robust phenotypic signals informative for therapy development (The Cancer Genome Atlas Research Network 2011; Verhaak et al. 2013; Maizels and Briscoe 2026), and a consensus signature without stratification remains highly informative for uncovering shared vulnerabilities and broad therapeutic opportunities across the ovarian cancer spectrum (Alexandrova et al. 2025; Cybula et al. 2021).

This study delineates the regulatory architecture that drives the transcriptional program of ovarian serous cystadenocarcinoma and prioritizes the master regulators most likely to sustain it. By reconstructing a mutual-information regulatory network over a robust cross-platform consensus signature and intersecting inferred regulatory activity with CRISPR-Cas9 dependency data, we identified 17 central transcriptional drivers that are both active in tumors and functionally essential for cancer-cell survival. These drivers converge on coherent oncogenic programs, mitotic and cell-cycle control, the DNA-damage checkpoint, chromatin remodeling, and apoptotic evasion, that together capture core hallmarks of high-grade serous disease. Several are established or emerging therapeutic targets, including *PLK1*, *CDK1*, *AURKB*, *CHEK1*, and survivin (*BIRC5*), and our results reinforce their relevance by recovering them through an orthogonal, dependency-aware strategy rather than from expression magnitude alone. Importantly, the analysis also surfaces transcriptional drivers that act through chromatin and RNA-processing mechanisms (*ACTL6A*, *SRSF2*) and a group of underexplored regulators (*KCMF1*, *GMNN*, *WDHD1*, *CHAF1B*, *ELOC*) that have not previously been linked to ovarian cancer dependency, nominating them as novel candidates for mechanistic and preclinical study. Beyond the specific biology, the integrative strategy itself offers a generalizable advantage: by combining multi-source meta-analysis with two complementary normal-control strategies, by extending Master Regulator Analysis beyond canonical transcription factors to the broader class of

transcriptional modulators, and by adding functional essentiality as a final filter, it overcomes the control ambiguity and reproducibility limitations that constrain conventional single-cohort approaches, and it is directly transferable to other malignancies where matched normal tissue is scarce. Collectively, these findings provide a prioritized, biologically interpretable map of the transcriptional drivers of ovarian cancer and a rational basis for selecting candidates, and combinations thereof, for experimental validation and therapeutic exploration.

Acknowledgements. Not applicable.

Funding. This research was funded by SECIHTI (CBF-2025-I-3339) to Martínez-Ramírez AS; Joel Ruiz-Hernández is a recipient of the fellowship (1100839), José Manuel Pastelín-Morales is a recipient of the fellowship (4066442).

Availability of data and materials. The data generated in the present study may be found in the Gene Expression Omnibus (GEO) database under accession numbers GSE14407, GSE18520, GSE27651, GSE38666, GSE40595 and GSE54388, at the following URL: https://www.ncbi.nlm.nih.gov/geo/ (accessed on July 1, 2026). The RNA-seq data generated in the present study may be found in The Cancer Genome Atlas (TCGA) at the following URL: https://portal.gdc.cancer.gov/, in the Genotype-Tissue Expression (GTEx) project at the following URL: https://gtexportal.org/, and in the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project at the following URL: https://ocg.cancer.gov/programs/target; all three were accessed through the OCTAD R package (v3.22) on July 1, 2026. All custom R scripts used for data processing, statistical analysis and figure generation in the present study may be found in the GitHub repository at the following URL: https://github.com/JoelRuiz26/Ovary_signatures.

Authors' contributions. Joel Ruiz-Hernández: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. Daniel Pérez-Calixto: Formal analysis, Investigation, Methodology, Supervision, Writing - review & editing. José Manuel Pastelín-Morales: Data curation, Investigation, Writing - review & editing. Enrique Hernández-Lemus: Investigation, Project administration, Supervision, Writing - review & editing. Genaro Vazquez-Victorio: Investigation, Supervision, Writing - review & editing. Angélica S. Martínez-Ramírez: Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Supervision, Writing - review & editing. Hugo Tovar: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing - original draft, Writing - review & editing. Joel Ruiz-Hernández and Hugo Tovar confirm the authenticity of all the raw data. All authors have read and approved the final manuscript.

Ethics approval and consent to participate. This study was conducted using publicly available, de-identified data from The Cancer Genome Atlas (TCGA/GDC), the Therapeutically Applicable Research to Generate Effective Treatments project (TARGET) and the Genotype-Tissue Expression project (GTEx), for which all ethical approvals and informed consent were obtained by the respective consortia in the

original studies. As the present study did not involve the recruitment of human participants or the collection of new human material, additional Institutional Review Board approval and informed consent were waived.

Patient consent for publication. Not applicable.

Competing interests. The authors declare that they have no competing interests.

Use of artificial intelligence tools. During the preparation of this manuscript, the authors used Claude (Opus 5; Anthropic, PBC, San Francisco, CA, USA) to assist with language editing and with selected R code implementation tasks. The authors reviewed and verified all such output and take full responsibility for the content of this publication.

Supplementary information. Supplementary Figures S1–S3 and Supplementary Tables SI–SII are provided as separate files.

Supplementary Table SI. Summary of control samples across anatomical sites derived from autoencoder approach. Control samples selected (normal and adjacent tissue) were the 88 most highly correlated with the ovarian tumor samples. Data derived from The Cancer Genome Atlas (TCGA) (The Cancer Genome Atlas Research Network 2011), the Genotype-Tissue Expression (GTEx) project (The GTEx Consortium 2020), and the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) project (Ma et al. 2018).

Supplementary Figure S1. Concordance between meta-analytic RRA signal and differential expression across independent cohorts.

Each panel shows the relationship between Robust Rank Aggregation (RRA) evidence and the DESeq2 differential-expression statistic for genes with robust signal (top $|RRA|$). The x-axis represents the magnitude and direction of the RRA signal as $\text{sign}(\text{RRA}) \cdot -\log_{10}(p_{\text{adj}})$, whereas the y-axis corresponds to the DESeq2 statistic for the TCGA *vs* AE contrast (left panel) or TCGA *vs* GTEx contrast (right panel). Red and blue points indicate genes classified as up-regulated or down-regulated according to RRA, respectively. Dashed lines mark the origin on both axes. Values reported in each panel indicate the Spearman correlation for genes with robust signal (ρtop) and for the full set of genes (ρglobal ranks). The most extreme genes (top 10 per direction) are explicitly labeled. The inset shows the relationship between absolute RRA and DESeq2 ranks on a logarithmic scale.

Supplementary Figure S2. Random subsampling confirms that the concordance between the RRA signal and DESeq2 statistics is not an artifact of subset size.

For each contrast (TCGA vs AE, left panel; TCGA vs GTEx, right panel), a null distribution was generated by randomly subsampling 1,000 sets of n = 2,000 genes from the full genome-wide pool. Histograms show the distribution of Spearman correlations obtained from those random subsets. The red dashed line indicates the observed correlation in the subset of the 2,000 genes with the strongest RRA evidence (AE: ρ = 0.546; GTEx: ρ = 0.590). In both contrasts, the observed correlation exceeds the entirety of the null distribution (empirical p ¡ 0.001), confirming that the improvement

in cross-platform concordance reflects a genuine property of the meta-analytic signal rather than a subsampling artifact.

Supplementary Figure S3. Expression variability across gene subsets defined by cohort overlap.

Violin plots show the distribution of the Fano factor (variance-to-mean ratio) for TCGA–AE and TCGA–GTEx cohorts across three nested gene subsets defined from differential expression results. **(A)** All expressed genes in each cohort (global distribution). **(B)** Differentially expressed genes shared between AE and GTEx (intersection of both RNA-seq cohorts). **(C)** Differentially expressed genes from each RNA-seq cohort that overlap with GEO-derived signals (AE∩GEO vs GTEx∩GEO). GEO data were used solely to define gene set membership and were not included in Fano factor calculations. Δ denotes the difference in median Fano factor between cohort distributions within each panel.

Supplementary Table SII. Literature evaluation of transcriptional drivers in ovarian cancer

Central transcriptional drivers from MRA that both modulate highly dependent core signature genes and are themselves essential in ovarian cancer cell lines (DepMap score $\leq$ -0.5). Canonical functions and key literature evidence are provided. (*****): also overexpressed in the core signature. (#): Transcriptional drivers that remain underexplored in ovarian cancer research, found in pan-cancer, general studies, predictions, or not ovary centered.

References

Acencio ML, Vazquez M, Chawla K, et al (2024) TFCheckpoint database update, a cross-referencing system for transcription factors from human, mouse and rat. Nucleic Acids Res 52:D334–D344

Affatato R, Carrassa L, Chilà R, et al (2020) Identification of PLK1 as a new therapeutic target in mucinous ovarian carcinoma. Cancers 12:672

Alexandrova E, Pecoraro G, Sellitto A, et al (2020) An overview of candidate therapeutic target genes in ovarian cancer. Cancers (Basel) 12:1470

Alexandrova E, Smal M, Di Rosa D, et al (2025) BRPF1 inhibition reduces migration and invasion of metastatic ovarian cancer cells, representing a potential therapeutic target. Sci Rep 15:7602

Alvarez MJ, Shen Y, Giorgi FM, et al (2016) Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 48:838–847

Amos SE, Choi YS (2021) The cancer microenvironment: Mechanical challenges of the metastatic cascade. Front Bioeng Biotechnol 9

Arafeh R, Shibue T, Dempster JM, et al (2025) The present and future of the cancer dependency map. Nat Rev Cancer 25:59–73

Auguste A, Blanc-Durand F, Deloger M, et al (2020) Small cell carcinoma of the ovary, hypercalcemic type (SCCOHT) beyond SMARCA4 mutations: A comprehensive genomic analysis. Cells 9

Behan FM, Iorio F, Picco G, et al (2019) Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. Nature 568:511–516

Benjamini Y, Drai D, Elmer G, et al (2001) Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125:279–284

Bray F, Laversanne M, Sung H, et al (2024) Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 74:229–263

Brown VE, Moore SL, Chen M, et al (2023) CDK2 regulates collapsed replication fork repair in CCNE1-amplified ovarian cancer cells via homologous recombination. NAR Cancer 5:zcad039

Califano A, Alvarez MJ (2017) The recurrent architecture of tumour initiation, progression and drug sensitivity. Nat Rev Cancer 17:116–130

Carro MS, Lim WK, Alvarez MJ, et al (2010) The transcriptional network for mesenchymal transformation of brain tumours. Nature 463:318–325

Caruso G, Weroha SJ, Cliby W (2025) Ovarian cancer: A review. JAMA 334:1278–1291

Chekalin E, Paithankar S, Zeng B, et al (2021) octad.db: Open cancer TherApeutic discovery (OCTAD) database. Bioconductor

Chen PM, Wong CN, Wong CN, et al (2023) Actin-like protein 6a expression correlates with cancer stem cell-like features and poor prognosis in ovarian cancer. Int J Mol Sci 24

Chen W, Zhuang X, Chen Y, et al (2025) Discovery of potent and selective CDK2 inhibitors with high safety and favorable bioavailability for the treatment of cancer. Eur J Med Chem 290:117503

Chen Y, Gao Y, Xu A, et al (2026) KCMF1 regulates HRI ubiquitination to inhibit the integrated stress response in ovarian cancer. Biochem Pharmacol 245:117634

Collado-Torres L, Nellore A, Kammers K, et al (2017) Reproducible RNA-seq analysis using recount2. Nat Biotechnol 35:319–321

Creighton CJ (2023) Gene expression profiles in cancers and their therapeutic implications. Cancer J 29:9–14

Cybula M, Wang L, Wang L, et al (2021) Patient-Derived xenografts of High-Grade serous ovarian cancer subtype as a powerful tool in Pre-Clinical research. Cancers

(Basel) 13:6288

Dang CV, Reddy EP, Shokat KM, et al (2017) Drugging the 'undruggable' cancer targets. Nat Rev Cancer 17:502–508

Davis S, Meltzer PS (2007) GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor. Bioinformatics 23:1846–1847

Dochez V, Caillon H, Vaucel E, et al (2019) Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. J Ovarian Res 12:28

Dolgalev I (2026) msigdbr: MSigDB gene sets for multiple organisms in a tidy data format. R package version 26.1.0, URL https://igordot.github.io/msigdbr/

Dou Z, Qiu C, Zhang X, et al (2022) HJURP promotes malignant progression and mediates sensitivity to cisplatin and WEE1-inhibitor in serous ovarian cancer. Int J Biol Sci 18:1188–1210

Edgar R, Domrachev M, Lash AE (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207–210

Fan Y, Pan Y, Jia L, et al (2024) BIRC5 facilitates cisplatin-chemoresistance in a m6A-dependent manner in ovarian cancer. Cancer Med 13:e6811

Ferreira PG, Muñoz-Aguirre M, Reverter F, et al (2018) The effects of death and post-mortem cold ischemia on human tissue transcriptomes. Nat Commun 9:490

Ganzfried BF, Riester M, Haibe-Kains B, et al (2013) curatedOvarianData: clinically annotated data for the ovarian cancer transcriptome. Database (Oxford) 2013:bat013

Gao T, Zhao P, Han S (2025a) Integrating bulk RNA-seq and scRNA-seq analyses with machine learning to predict platinum response and prognosis in ovarian cancer. Sci Rep 15:19123

Gao Y, Zheng K, Tan H, et al (2025b) Single-cell RNA sequencing reveals the intratumoral heterogeneity and immune microenvironment of small cell carcinoma of the ovary, hypercalcemic type. J Ovarian Res 18:76

Garlisi B, Lauks S, Aitken C, et al (2024) The complex tumor microenvironment in ovarian cancer: Therapeutic challenges and opportunities. Curr Oncol 31:3826–3844

Gasimli K, Raab M, Tahmasbi Rad M, et al (2022) Sequential targeting of PLK1 and PARP1 reverses the resistance to PARP inhibitors and enhances Platin-Based chemotherapy in BRCA-Deficient High-Grade serous ovarian cancer with KRAS amplification. Int J Mol Sci 23:10892

Geistlinger L, Oh S, Ramos M, et al (2020) Multi-omic analysis of subtype evolution and heterogeneity in high-grade serous ovarian carcinoma. Cancer Res 80:4335–4345

Glaab E, Rauschenberger A, Banzi R, et al (2021) Biomarker discovery studies for patient stratification using machine learning analysis of omics data: a scoping review. BMJ Open 11:e053674

Gui T, Yao C, Jia B, et al (2021) Identification and analysis of genes associated with epithelial ovarian cancer by integrated bioinformatics methods. PloS One 16:e0253136

Guven E (2023) Decision of the optimal rank of a nonnegative matrix factorization model for gene expression data sets utilizing the unit invariant knee method: Development and evaluation of the elbow method for rank selection. JMIR Bioinforma Biotechnol 4:e43665

Han J, Xie R, Yang Y, et al (2021) CENPA is one of the potential key genes associated with the proliferation and prognosis of ovarian cancer based on integrated bioinformatics analysis and regulated by MYBL2. Transl Cancer Res 10:4076–4086

Hart T, Chandrashekhar M, Aregger M, et al (2015) High-Resolution CRISPR screens reveal fitness genes and Genotype-Specific cancer liabilities. Cell 163:1515–1526

Huang Y, Fan Y, Zhao Z, et al (2023) Inhibition of CDK1 by RO-3306 exhibits Anti-Tumorigenic effects in ovarian cancer cells and a transgenic mouse model of ovarian cancer. Int J Mol Sci 24

Jaladanki SK, Elmas A, Malave GS, et al (2021) Genetic dependency of alzheimer's disease-associated genes across cells and tissue types. Sci Rep 11:12107

Jeong S, Jo YS, Park S, et al (2025) Aggressive serous ovarian cancer subtype defined by high centrality lncRNA profiles and master transcription factors. Sci Rep 15:1–15

Kerslake R, Belay B, Panfilov S, et al (2023) Transcriptional landscape of 3d vs. 2d ovarian cancer cell models. Cancers (Basel) 15:3350

Kolberg L, Raudvere U, Kuzmin I, et al (2020) gprofiler2 – an r package for gene list functional enrichment analysis and namespace conversion toolset g:profiler. F1000Research 9:ELIXIR–709

Kolde R, Laur S, Adler P, et al (2012) Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics 28:573–580

Lachmann A, Giorgi FM, Lopez G, et al (2016) ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 32:2233–2235

Lambert SA, Jolma A, Campitelli LF, et al (2018) The human transcription factors. Cell 175:598–599

Lashen A, Algethami M, Alqahtani S, et al (2024) The clinicopathological significance of the cyclin D1/E1–Cyclin-Dependent kinase (CDK2/4/6)–retinoblastoma (RB1/pRB1) pathway in epithelial ovarian cancers. Int J Mol Sci 25

Lazar C, Meganck S, Taminau J, et al (2013) Batch effect removal methods for microarray gene expression data integration: a survey. Brief Bioinform 14:469–490

Lefebvre C, Rajbhandari P, Alvarez MJ, et al (2010) A human b-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol Syst Biol 6:377

Li F, Zhao C, Diao Y, et al (2022) MEX3A promotes the malignant progression of ovarian cancer by regulating intron retention in TIMELESS. Cell Death Dis 13:553

Li Y, Chen Z, Peng J, et al (2023) The splicing factor SNRPB promotes ovarian cancer progression through regulating aberrant exon skipping of POLA1 and BRCA2. Oncogene 42:2386–2401

Liberzon A, Birger C, Thorvaldsdóttir H, et al (2015) The molecular signatures database hallmark gene set collection. Cell Syst 1:417–425

Liu B, Fan Y, Lv C, et al (2025a) Circular RNA circMYBL2 regulates the progression of ovarian cancer through miR-195-5P/BIRC5 axis. J Ovarian Res 19:40

Liu Y, Liu H, Zhu C, et al (2025b) Tumor small extracellular vesicle-transmitted LncRNA CATED promotes platinum-resistance in high-grade serous ovarian cancer. Adv Sci (Weinh) 12:e05963

Lliberos C, Richardson G, Papa A (2024) Oncogenic pathways and targeted therapies in ovarian cancer. Biomolecules 14:585

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550

Ma H, Qi G, Han F, et al (2022) HMGB3 promotes PARP inhibitor resistance through interacting with PARP1 in ovarian cancer. Cell Death Dis 13:263

Ma H, Qi G, Han F, et al (2023) HMGB3 promotes the malignant phenotypes and stemness of epithelial ovarian cancer through the MAPK/ERK signaling pathway. Cell Commun Signal CCS 21:144

Ma X, Liu Y, Liu Y, et al (2018) Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature 555:371–376

Maizels RJ, Briscoe J (2026) Gene regulatory networks: from correlative models to causal explanations. Nat Rev Genet https://doi.org/10.1038/s41576-026-00939-1

Meyers RM, Bryan JG, McFarland JM, et al (2017) Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. Nat Genet 49:1779–1784

Michiels S, Ternès N, Rotolo F (2016) Statistical controversies in clinical research: prognostic gene signatures are not (yet) useful in clinical practice. Ann Oncol 27:1–9

Micoli G, Lavikka K, Li Y, et al (2025) Decoding the genomic and functional landscape of emerging subtypes in ovarian cancer. Cancer Discov 15:2262–2277

Millstein J, Budden T, Goode EL, et al (2020) Prognostic gene expression signature for high-grade serous ovarian cancer. Ann Oncol 31:1240–1250

Mirabelli P, Coppola L, Salvatore M (2019) Cancer cell lines are useful model systems for medical research. Cancers 11:1098

Moskovich D, Beilinson D, Rosemarin A, et al (2025) DIO3 depletion attenuates ovarian cancer growth via reduced glycolysis and alterations in glutamine metabolism. Mol Metab 100:102225

Moss EL, Hollingworth J, Reynolds TM (2005) The role of CA125 in clinical practice. J Clin Pathol 58:308–312

Nair J, Huang TT, Murai J, et al (2020) Resistance to the CHK1 inhibitor prexasertib involves functionally distinct CHK1 activities in BRCA wild-type ovarian cancer. Oncogene 39:5520–5535

Otto T, Sicinski P (2017) Cell cycle proteins as promising targets in cancer therapy. Nat Rev Cancer 17:93–115

Padi M, Quackenbush J (2015) Integrating transcriptional and protein interaction networks to prioritize condition-specific master regulators. BMC Syst Biol 9:80

Parmar K, Kochupurakkal BS, Lazaro JB, et al (2019) The CHK1 inhibitor prexasertib exhibits monotherapy activity in High-Grade serous ovarian cancer models and sensitizes to PARP inhibition. Clin Cancer Res 25:6127–6140

Pellarin I, Dall'Acqua A, Gambelli A, et al (2020) Splicing factor proline- and glutamine-rich (SFPQ) protein regulates platinum response in ovarian cancer-modulating SRSF2 activity. Oncogene 39:4390–4403

Pérez-Fidalgo JA, Gambardella V, Pineda B, et al (2020) Aurora kinases in ovarian cancer. ESMO Open 5:e000718

Risso D, Ngai J, Speed TP, et al (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol 32:896–902

Ritchie ME, Phipson B, Wu D, et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47

Ruiz-Hernández J, de Anda-Jáuregui G, Hernández-Lemus E (2025) Drug repositioning for HPV Clade-Specific cervicouterine cancer using the OCTAD pipeline. Int J Mol Sci 26:11238

Schofield JA, Hahn S (2024) Transcriptional noise, gene activation, and roles of SAGA and mediator tail measured using nucleotide recoding single-cell RNA-seq. Cell Rep 43:114593

Sehgal N (2019) Efficacy of color doppler ultrasonography in differentiation of ovarian masses. J Midlife Health 10:22–28

Sereti K, Russo AE, Raisner R, et al (2025) PAX8 interacts with the SWI/SNF complex at enhancers to drive proliferation in ovarian cancer. Mol Cancer Res 23:416–425

Shen F, Li C, Li Y, et al (2025a) RNA sequencing in ovarian cancer research: a comprehensive review. J Ovarian Res 18:308

Shen J, Yu S, Sun X, et al (2019) Identification of key biomarkers associated with development and prognosis in patients with ovarian carcinoma: evidence from bioinformatic analysis. J Ovarian Res 12:110

Shen X, Wang J, Chen S, et al (2025b) Clinical relevance of PLK1 in epithelial ovarian cancer. Front Pharmacol 16

Shi X, Gekas C, Verduzco D, et al (2024) Building a translational cancer dependency map for the cancer genome atlas. Nat Cancer 5:1176–1194

Sine CC, Watts LP, Fernandez B, et al (2025) p16 expression confers sensitivity to CDK2 inhibitors in cyclin e1–driven ovarian cancers. Sci Signal 18:eadv0415

Singh A, Gupta S, Badarukhiya JA, et al (2020) Detection of aberrant methylation of HOXA9 and HIC1 through multiplex MethyLight assay in serum DNA for the early detection of epithelial ovarian cancer. Int J Cancer 147:1740–1752

Smith HL, Willmore E, Prendergast L, et al (2024) ATR, CHK1 and WEE1 inhibitors cause homologous recombination repair deficiency to induce synthetic lethality with PARP inhibitors. Br J Cancer 131:905–917

Sykaras AG, Pergaris A, Theocharis S (2021) Challenging, accurate and feasible: CAF-1 as a tumour proliferation marker of diagnostic and prognostic value. Cancers 13:2575

Tan R, Wen M, Yang W, et al (2025a) Integrated proteomics and scRNA-seq analyses of ovarian cancer reveal molecular subtype-associated cell landscapes and immunotherapy targets. Br J Cancer 132:111–125

Tan XQ, Guo AY, Zheng LF, et al (2025b) Research progress on FOXM1 in ovarian cancer diagnosis and therapeutics. Front Oncol 15:1598868

Taş M, Utkan Korun ZE, Koçak A, et al (2019) Comparison of adenosine deaminase (ADA) and cancer antigen 125 (CA 125) levels in epithelial ovarian tumors. HAYDARPAŞA NUMUNE Med J 59:309–314

The Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. Nature 474:609–615

The GTEx Consortium (2020) The GTEx consortium atlas of genetic regulatory effects across human tissues. Science 369:1318–1330

Tovar H, García-Herrera R, Espinal-Enríquez J, et al (2015) Transcriptional master regulator analysis in breast cancer genetic networks. Comput Biol Chem 59:67–77

Tsherniak A, Vazquez F, Montgomery PG, et al (2017) Defining a cancer dependency map. Cell 170:564–576.e16

Tudose C, Bond J, Ryan CJ (2023) Gene essentiality in cancer is better predicted by mRNA abundance than by gene regulatory network-inferred activity. NAR Cancer 5:zcad056

Vaicekauskaitė I, Kazlauskaitė P, Gineikaitė R, et al (2025) Integrative analysis of gene expression and promoter methylation to differentiate High-Grade serous ovarian cancer from benign tumors. Biomedicines 13

Verhaak RG, Tamayo P, Yang JY, et al (2013) Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest 123:517–525

Vivian J, Rao AA, Nothaft FA, et al (2017) Toil enables reproducible, open source, big biomedical data analyses. Nat Biotechnol 35:314–316

Võsa U, Kolde R, Vilo J, et al (2014) Comprehensive meta-analysis of microRNA expression using a robust rank aggregation approach. Methods Mol Biol 1182:361–373

Waldron L, Haibe-Kains B, Culhane AC, et al (2014) Comparative meta-analysis of prognostic gene signatures for late-stage ovarian cancer. J Natl Cancer Inst 106:dju049

Wang B, Li X, Zhao G, et al (2018a) mir-203 inhibits ovarian tumor metastasis by targeting BIRC5 and attenuating the TGFβ pathway. J Exp Clin Cancer Res 37:235

Wang H, Fang L, Jiang J, et al (2018b) The cisplatin-induced lncRNA PANDAR dictates the chemoresistance of ovarian cancer via regulating SFRS2-mediated p53 phosphorylation. Cell Death Dis 9:1103

Wang J, Hu M, Min J, et al (2024) A positive feedback loop of SRSF9/USP22/ZEB1 promotes the progression of ovarian cancer. Cancer Biol Ther 25:2427415

Wang JM, Liu N, Wei XJ, et al (2026) Correction to: Regulation of AUF1 alternative splicing by hnRNPA1 and SRSF2 modulate the sensitivity of ovarian cancer cells to cisplatin. Cell Oncol 49:17

Wang JZ, Du Z, Payattakool R, et al (2007) A new method to measure the semantic similarity of GO terms. Bioinformatics 23:1274–1281

Wang L, Zhang Q, Wang X, et al (2025) Therapeutic landscape of ovarian cancer: recent advances and emerging therapies. Biomark Res 13:103

Weir A, Kang EY, Meagher NS, et al (2023) Increased FOXJ1 protein expression is associated with improved overall survival in high-grade serous ovarian carcinoma: an ovarian tumor tissue analysis consortium study. Br J Cancer 128:137–147

Wen Y, Hou Y, Yi X, et al (2021) EZH2 activates CHK1 signaling to promote ovarian cancer chemoresistance by maintaining the properties of cancer stem cells. Theranostics 11:1795–1813

Willis S, Villalobos VM, Gevaert O, et al (2016) Single gene prognostic biomarkers in ovarian cancer: A Meta-Analysis. PloS One 11:e0149183

Xiao Y, Lin FT, Lin WC (2021) ACTL6A promotes repair of cisplatin-induced DNA damage, a new mechanism of platinum resistance in cancer. Proc Natl Acad Sci 118:e2015808118

Xing X, Gu F, Hua L, et al (2021) TIMELESS promotes tumor progression by enhancing macrophages recruitment in ovarian cancer. Front Oncol 11

Xu C, Chen C, Xu Y, et al (2025) Prognostic significance of CDK1 in ovarian and cervical cancers. J Cancer 16:1656–1667

Xu X, Ouyang L, Wang J, et al (2026) KCMF1 promotes malignant progression by NXN ubiquitin-dependent degradation in ovarian cancer. Cell Cycle 25:1–19

Yang W, Cho H, Shin HY, et al (2016) Accumulation of cytoplasmic cdk1 is associated with cancer growth and survival rate in epithelial ovarian cancer. Oncotarget 7:49481–49497

Yu G, Wang LG, Han Y, et al (2012) clusterprofiler: an r package for comparing biological themes among gene clusters. OMICS J Integr Biol 16:284–287

Zeng B, Glicksberg BS, Newbury P, et al (2021) OCTAD: an open workspace for virtually screening therapeutics targeting precise cancer patient groups using gene expression features. Nat Protoc 16:728–753

Zeng WZD, Glicksberg BS, Li Y, et al (2019) Selecting precise reference normal tissue samples for cancer research using a deep learning approach. BMC Med Genomics 12:21

Zhang C, Ni X, Tao C, et al (2024a) Targeting PUF60 prevents tumor progression by retarding mRNA decay of oxidative phosphorylation in ovarian cancer. Cell Oncol (Dordr) 47:157–174

Zhang J, Zhang J, Wei Y, et al (2019) ACTL6A regulates follicle-stimulating hormone-driven glycolysis in ovarian cancer cells via PGK1. Cell Death Dis 10:811

Zhang Y, Zhang W, Sun L, et al (2022) HJURP inhibits proliferation of ovarian cancer cells by regulating CENP-A/CENP-N. Bull Cancer 109:1007–1016

Zhang Y, Wu Y, Shi X, et al (2024b) m6A-mediated hsa_circ_0061179 inhibits DNA damage in ovarian cancer cells via miR-143-3p/TIMELESS. Mol Carcinog 63:1542–1558

Zhang Z, Zhu Q (2023) WD repeat and HMG box DNA binding protein 1: An oncoprotein at the hub of tumorigenesis and a novel therapeutic target. Int J Mol Sci 24

Zhao G, Newbury P, Ishi Y, et al (2022) Reversal of cancer gene expression identifies repurposed drugs for diffuse intrinsic pontine glioma. Acta Neuropathol Commun 10:150

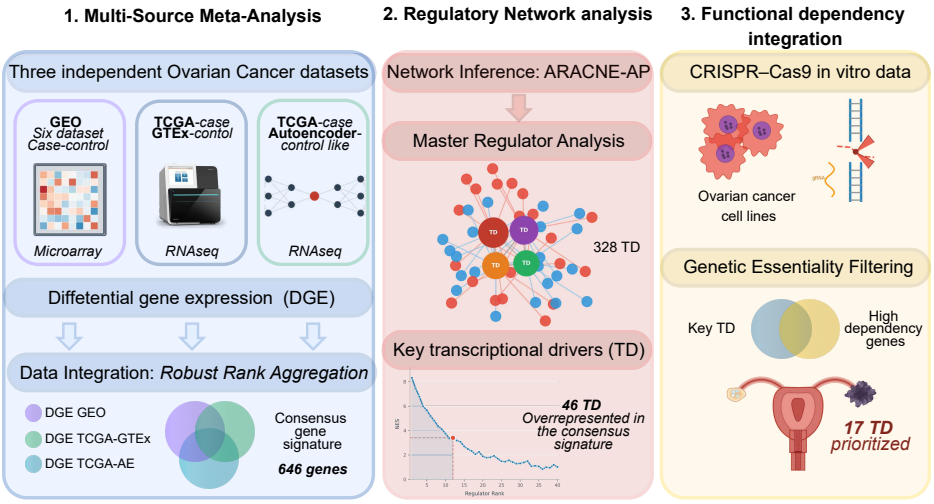


Fig. 1 Methodological pipeline for integrative identification of functionally essential transcriptional drivers in ovarian cancer. The workflow comprises three sequential analytical layers: **(1)** Multi-source transcriptomic analysis: six GEO microarray datasets were preprocessed independently, analyzed for differential expression with limma, and integrated by Robust Rank Aggregation (RRA) to generate a directional gene list. In parallel, TCGA ovarian RNA-seq samples were compared against two alternative control cohorts, GTEx normal ovarian tissue and autoencoder-selected normal-like samples, using DESeq2. Differentially expressed genes from the three independent sources were integrated by requiring consistent directionality (up- or downregulation) across all contrasts, yielding a consensus gene signature. **(2)** Regulatory network inference and Master Regulator Analysis: an ovarian cancer-specific regulatory network was inferred from TCGA tumor expression profiles using ARACNe-AP and converted into regulons. msVIPER was then applied to evaluate enrichment of the consensus signature within the transcriptional neighborhoods of an expanded regulator space defined by TFcheckpoint 2, including transcription factors, cofactors, chromatin remodelers, and epigenetic regulators. Regulators were ranked according to normalized enrichment scores and statistical significance to identify high-confidence key transcriptional drivers **(3)** Functional dependency integration: CRISPR–Cas9 dependency data from the Cancer Dependency Map (DepMap) were used to assess gene essentiality across ovarian cancer cell lines. Regulators were mapped to median dependency scores, and those meeting the essentiality threshold were intersected with the high-confidence regulatory candidates to prioritize functionally essential central transcriptional drivers for downstream experimental validation.

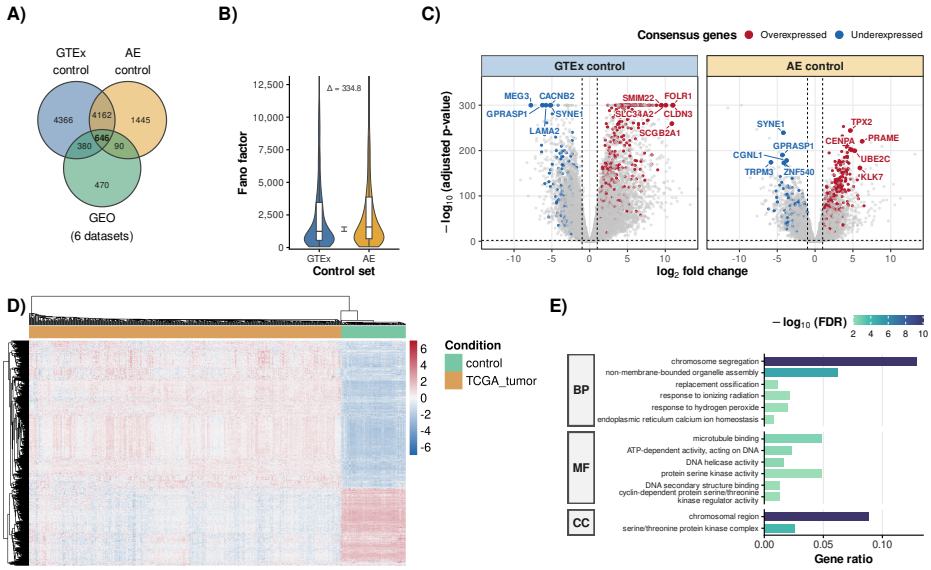


Fig. 2 Consensus gene expression signature across ovarian cancer transcriptomic datasets. (A) Venn diagram showing the overlap of significant genes identified in three independent contrasts: ovarian tumors vs normal ovarian tissues (TCGA-OV vs GTEx), ovarian tumors vs normal-like controls selected via an autoencoder-based approach (TCGA-OV vs AE), and the meta-analysis of six microarray studies (GEO). The three-way intersection defines the consensus gene set. **(B)** Distribution of Fano factor values for the consensus gene set, computed from RNA-seq data in TCGA-OV tumor samples normalized against GTEx and AE controls; Δ denotes median differences between groups. **(C)** Volcano plots of differential expression for TCGA-OV vs GTEx (top) and TCGA-OV vs AE (bottom), showing log2 fold change versus –log10 adjusted p-value. Genes from the consensus set are highlighted according to direction of regulation (red, overexpressed; blue, underexpressed), with top-ranked genes annotated. **(D)** Heatmap of the consensus gene set (rows = genes, columns = tumor samples). Expression values are variance-stabilizing transformed (VST) and row-scaled (Z-score). Hierarchical clustering groups samples and genes, illustrating the coherent signature's expression pattern across tumors. **(E)** Over-representation analysis (ORA) of the consensus gene set showing representative enriched biological processes (bar length = gene count; color = –log10 adjusted p-value).

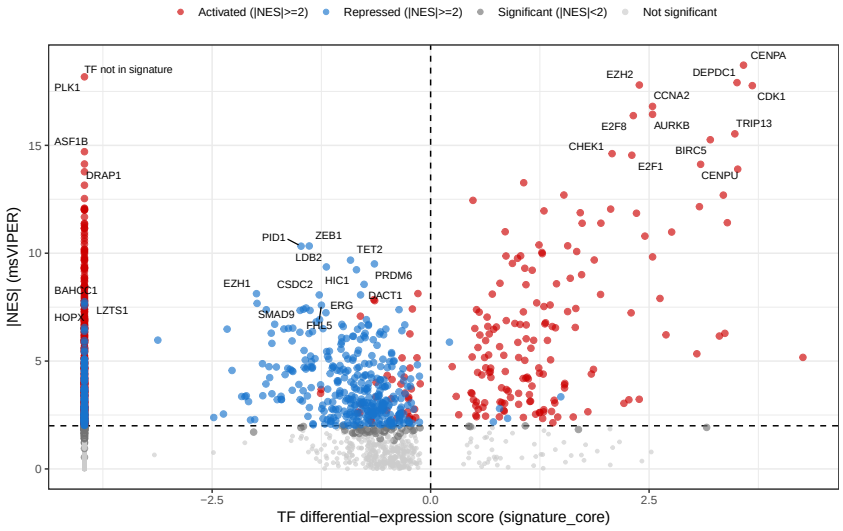


Fig. 3 Relationship between key transcriptional drivers differential expression and inferred regulatory activity from MRA using the core signature. The x-axis shows the transcription modulators differential-expression score derived from the core signature (signature_core), whereas the y-axis shows the absolute value of the msVIPER Normalized Enrichment Score ($|NES|$) as a measure of regulatory activity magnitude. Points are colored by regulatory status: activated regulators (NES $\geq$ 2, red), repressed regulators (NES $\leq$ –2, blue), significant weak regulators (empirical FDR ¡ 0.05 and $|NES|$ ¡ 2, dark gray), and non-significant regulators (light gray). Dashed lines indicate the activity and expression thresholds ($|NES| = 2$ and signature_core = 0). Transcription modulators not present in the core signature are shown as a separate column on the left of the x-axis. Labels highlight the 15 activated and 15 repressed regulators with the largest $|NES|$ among those with empirical FDR ¡ 0.05.

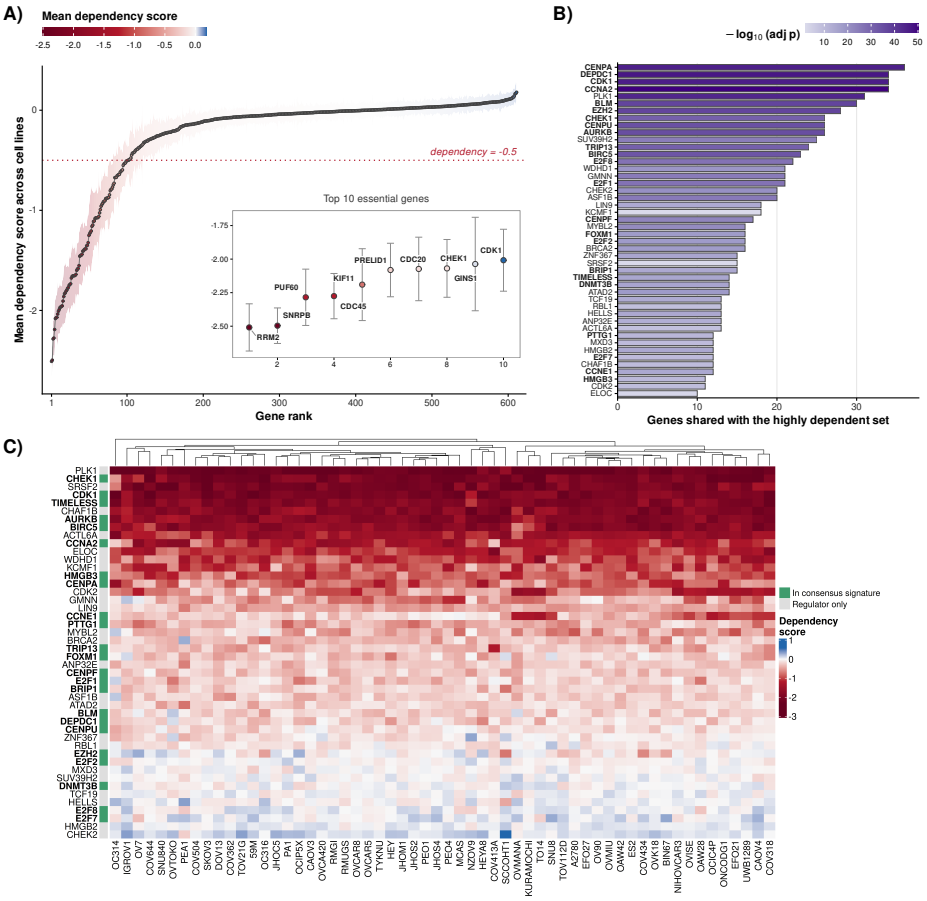


Fig. 4 Transcriptional drivers associated with highly dependent genes in ovarian cancer cell lines. (A) Ranked distribution of median gene CRISPR–Cas9 dependency scores from DepMap for the core signature (n = 646 genes), where more negative values indicate greater essentiality. Highly dependent genes (HD-genes), defined by a dependency score $\leq$ –0.5, are indicated by the dashed red line. (B) Over-representation analysis of his transcriptional neighborhood using HD-genes. Transcriptional drivers (TDs) are ranked by the number of overlapping neighbor genes. Color scale represents statistical significance (–log10 p-value). (C) Heatmap of dependency scores for enriched master regulators across cell lines. Transcription modulators are ordered by median dependency, with more negative values indicating higher essentiality. Bold labels indicate transcriptional modulators that are also differentially expressed.

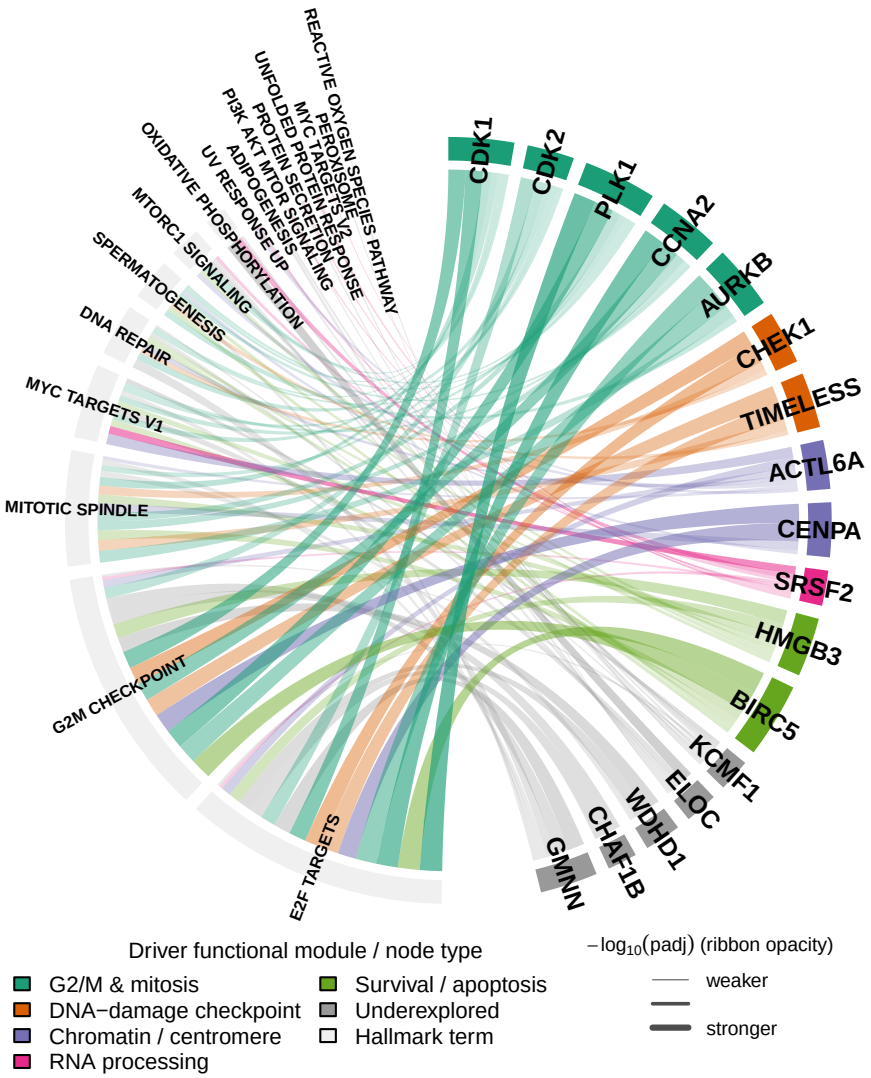


Fig. 5 Bipartite chord diagram linking the 17 prioritized transcriptional drivers to Hallmark gene sets (MSigDB) significantly enriched within their ARACNe-AP regulons. Arcs on the right represent the 17 drivers, colored by functional module (G2/M and mitosis, DNA-damage checkpoint, chromatin/centromere, RNA processing, survival/apoptosis, and underexplored); arcs on the left represent Hallmark terms, shown in neutral gray to distinguish this node type. Arc width is proportional to GeneRatio (the percentage of each driver's regulon annotated to a given term, summed over all of its significant terms), so drivers with large regulons but modest relative enrichment (e.g., KCMF1) appear with narrower arcs than drivers with concentrated enrichment in few programs (e.g., PLK1, CDK1, CCNA2, AURKB, BIRC5); ribbon opacity is proportional to -log10(adjusted p-value) of the enrichment. Of 533 driver-term pairs tested, 99 passed the significance filter (adjusted p¡0.05, Benjamini-Hochberg), dominated by E2F TARGETS (16/17 drivers), G2M CHECKPOINT (16/17), and MITOTIC SPINDLE (13/17).

Supplementary Files

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

- OvarianFIGSupplementaryFile.zip