

Circulating Long Non-Coding RNAs as Superior Non-Invasive Biomarkers for Prostate Cancer Stratification: Integrative Diagnostic Evaluation and Systems-Level Oncogenic Network Analysis

Poornima Verma¹

poornimaverma555@gmail.com

Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS)

shivani kumari

Sanjay Gandhi Post Graduate Institute of Medical Sciences

Research Article

Keywords: Prostate cancer, liquid biosies, biomarker

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11267022/v1>

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

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Prostate cancer (PCa) remains one of the leading causes of cancer-related morbidity and mortality among men worldwide, largely due to limitations in early detection and reliable stratification of aggressive disease. Although serum prostate-specific antigen (PSA) is widely used for screening, its limited specificity necessitates the identification of more robust non-invasive biomarkers. Long non-coding RNAs (lncRNAs) have emerged as critical regulators of tumorigenesis and metastasis and hold promise as circulating biomarkers. The present study aimed to evaluate the diagnostic and prognostic potential of selected circulating lncRNAs (HOTAIR, MALAT1, SCHLAP1, PCAT, GAS5, FR0348383, and lincRNA-p21) in prostate cancer. Quantitative expression profiling was performed, followed by receiver operating characteristic (ROC) analysis to assess discriminatory performance relative to PSA and Gleason score. In parallel, integrative bioinformatics analyses including protein–protein interaction (PPI) networking, Gene Ontology enrichment, and pathway mapping were conducted to elucidate underlying molecular mechanisms. Our findings demonstrate significant upregulation of HOTAIR, MALAT1, SCHLAP1, PCAT, and lincRNA-p21 in metastatic disease, with MALAT1 and SCHLAP1 showing superior diagnostic accuracy (AUC up to ~ 0.98) compared with PSA. Network analysis revealed central oncogenic hubs including MTOR, AKT1, TP53, EGFR, and MYC, highlighting coordinated regulation of epigenetic remodeling, metabolic reprogramming, and proliferative signaling. Functional enrichment further indicated strong involvement of hypoxia response, EMT, and Wnt signaling pathways. Collectively, this study supports circulating lncRNAs as powerful non-invasive biomarkers for prostate cancer stratification and provides mechanistic insight into the integrated oncogenic–metabolic network driving disease progression. These findings may facilitate improved risk assessment and precision management of prostate cancer.

1. Introduction

Approximately 1.4 million men worldwide receive a prostate cancer (PCa) diagnosis annually, yet the fundamental clinical imperative—distinguishing lethal disease from indolent tumours that require only surveillance—remains largely unmet [Global Cancer Observatory (GLOBOCAN), 2020]. The dominant screening tool, serum prostate-specific antigen (PSA), achieves a positive predictive value of only 25–40% within the diagnostic grey zone (PSA 4–10 ng/mL), producing an epidemic of unnecessary biopsies and overtreatment while simultaneously failing to reliably identify the aggressive phenotypes responsible for PCa mortality (United States Preventive Services Task Force, 2018). Improving this situation demands not a more sensitive test for cancer presence, but a molecularly grounded biomarker that reads tumour biology directly—one capable of discriminating between watchful-waiting candidates and patients requiring immediate systemic intervention [1, 2].

Long non-coding RNAs (lncRNAs)—transcripts exceeding 200 nucleotides without protein-coding potential—have emerged as precisely such candidates. These molecules regulate gene expression at epigenetic, transcriptional, and post-transcriptional levels, and their dysregulation underpins virtually all hallmarks of cancer including proliferation, EMT, metabolic reprogramming, immune evasion, and

metastasis (ENCODE Project Consortium, 2012; Mattick et al., 2023 [5]). Crucially, lncRNAs are detectable in body fluids such as serum and urine in a cell-free form, rendering them practical liquid biopsy targets (National Cancer Institute, 2020).

Among the best-characterised oncogenic lncRNAs in PCa, HOTAIR orchestrates chromatin remodelling through interaction with Polycomb repressive complex 2, promoting epigenetic silencing of tumour suppressor genes [1]. MALAT1 modulates alternative splicing, androgen receptor signalling, and has been validated as a urinary biomarker for PCa versus BPH in prospective multicentre cohorts [2, 15]. SCHLAP1 antagonises the SWI/SNF chromatin-remodelling complex and is robustly linked to aggressive disease phenotypes and lymph node metastasis [3, 6,7]. Additional candidates—PCAT-1, GAS5, lincRNA-p21, and FR0348383—have been reported to influence tumour biology at the tissue level, although their combined serum diagnostic performance in distinct clinical subgroups, including metastatic disease, has received comparatively limited attention [4, 5].

Despite this prior work, important gaps persist. Most published lncRNA biomarker studies focus on tissue expression or urinary lncRNAs in predominantly localised disease cohorts; simultaneous serum profiling of a multi-marker lncRNA panel across clinically stratified PCa subgroups—particularly metastatic disease where novel biomarkers are most urgently needed—remains incompletely explored. No published study has simultaneously characterised the serum expression of this specific seven-marker panel alongside integrative systems-level network analysis of the molecular programmes these lncRNAs are embedded in [8,9]. Furthermore, PCa progression is increasingly understood as a systems-level process involving coordinated dysregulation of oncogenic signalling, metabolic adaptation, and epigenetic remodelling (The Cancer Genome Atlas Research Network, 2015 [16,17]. Network-based approaches provide mechanistic depth that individual biomarker studies alone cannot achieve. The present pilot study uniquely addresses this gap by: (i) simultaneously quantifying serum expression of seven lncRNAs across clinically stratified PCa patients and healthy controls; (ii) applying rigorous ROC analysis with confidence interval estimation; and (iii) contextualising the identified biomarker signals within a literature-curated gene interaction network to elucidate their mechanistic underpinnings and guide hypothesis generation for prospective validation.

2.MATERIALS AND METHODS

2.1 Study Design and Patient Cohort

This pilot case–control study included clinically characterised PCa patients and age-matched healthy controls. Patients were stratified into clinical subgroups—overall PCa, high-risk disease, and metastatic PCa—based on standard clinicopathological criteria including PSA level, Gleason score, and imaging findings. Written informed consent was obtained from all participants, and the study protocol was approved by the Institutional Ethics Committee in accordance with the Declaration of Helsinki. Sample Size and Power Calculation. Given the exploratory pilot nature of this study, a formal a priori power analysis was conducted to characterise the statistical properties of the cohort. Assuming an expected

AUC of 0.90, a null AUC of 0.50, $\alpha = 0.05$ (two-tailed), and power = 0.80, a minimum of $n = 38$ patients per group would be required for a single-biomarker ROC study. The current pilot cohort of $n = 10$ per group provides approximately 32% power to detect AUC = 0.90 and approximately 85% power to detect AUC ≥ 0.98 ; however, AUC estimates from small pilot samples are frequently inflated relative to true population values due to overfitting. Accordingly, all findings are presented as preliminary, hypothesis-generating estimates requiring prospective validation in adequately powered cohorts (minimum $n = 80-100$ per group for single-centre studies).

Inclusion Criteria:

- Histopathologically confirmed prostate cancer
- Availability of complete clinical data
- Signed informed consent

Exclusion Criteria:

- History of other malignancies
- Active inflammatory or infectious disease
- Inadequate sample quality or haemolysed serum

Note on prior treatment: Patients who had received prior chemotherapy (4/10, 40% of the metastatic group) or androgen deprivation therapy (ADT) were retained in the

analysis, as treatment status was recorded for all participants. Chemotherapy exposure was included as a covariate in supplementary multivariable logistic regression models. Given the pilot sample size, no patients were excluded on the basis of prior systemic therapy in order to preserve statistical power; this represents a recognised limitation that future prospective studies must address using stratified enrolment with pre-specified criteria for prior-treatment status.

2.2 Sample Collection and Processing

Pre-analytical variables. Peripheral blood samples were collected in serum separation tubes (SST with clot activator) and allowed to coagulate at room temperature for 30 minutes. Serum was separated by centrifugation at 3,000 rpm (approximately $1,000\times g$) for 10 minutes at 4°C within 2 hours of venipuncture. All samples were aliquoted (200 μL

per aliquot) and stored at -80°C . The maximum number of freeze–thaw cycles was two. Haemolysed samples were excluded; haemolysis was assessed visually and confirmed by spectrophotometric absorbance at 414 nm. Pre-analytical variables recorded for each sample included: collection tube type, time from venipuncture to centrifugation, storage duration, freeze–thaw cycle count, and haemolysis score. To control for RNA isolation variability across samples, a synthetic *Caenorhabditis elegans* miRNA spike-in control (cel-miR-39; 1.6×10^8 copies per mL serum) was added to each sample immediately after thawing and prior to RNA extraction.

2.3 RNA Isolation and cDNA Synthesis

Total RNA enriched for circulating non-coding RNAs was isolated from 400 µL serum per sample using a commercially validated phenol-free extraction protocol (miRNeasy Serum/Plasma Kit, Qiagen) according to the manufacturer's instructions. RNA yield was quantified using a Qubit RNA High Sensitivity assay (Thermo Fisher Scientific). Size distribution and RNA quality were assessed on an Agilent 2100 Bioanalyzer using the small RNA chip to confirm recovery of RNA fragments in the appropriate size range for circulating lncRNA species; this approach complies with MIQE guidelines for liquid biopsy RNA assessment (Bustin et al., 2009, Clinical Chemistry). The A260/A280 ratio was recorded as a supplementary purity measure. Samples falling below a minimum yield of 10 ng total RNA per mL serum or showing evidence of severe degradation were excluded from further analysis. Complementary DNA (cDNA) was synthesised from 1 µg total RNA using a high-capacity reverse transcription kit with random hexamer priming (Applied Biosystems). All reactions included no-template controls (NTC) and no-reverse transcriptase controls (no-RT) to detect contamination and genomic DNA carry-over, respectively.

2.4 Quantitative Real-Time PCR (qRT-PCR)

Expression levels of seven lncRNAs—HOTAIR, MALAT1, SCHLAP1, PCAT-1 (also designated PCAT1; HGNC:37514; ENSEMBL: ENSG00000230061), GAS5, lincRNA-p21, and FR0348383 (FANTOM5 CAGEseq catalogue identifier; this lncRNA represents a putative intergenic transcript provisionally mapped to chromosome 14q based on FANTOM5 atlas annotations; a formal HGNC-approved gene symbol has not been assigned at the time of this study; its selection was based on preliminary evidence of altered expression in prostate cancer transcriptomic datasets)—were quantified by SYBR Green-based qRT-PCR on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad).

PCR conditions:

- Initial denaturation: 95°C for 5 min
- 40 amplification cycles: 95°C for 15 s; 60°C for 60 s
- Melt curve analysis: 65–95°C in 0.5°C increments to confirm single amplicon specificity

Primer sequences, amplicon sizes, GenBank accession numbers, and target isoform information for all seven lncRNAs and the endogenous reference gene are provided in Supplementary Table S2. All primer pairs were validated for single-amplicon specificity by melt curve analysis and agarose gel electrophoresis. PCR efficiencies were calculated from five-point serial dilution standard curves and fell within the acceptable range of 92–108% for all assays, in full compliance with MIQE guidelines (Bustin et al., 2009).

Relative expression was calculated using the $2^{-\Delta\Delta C_t}$ method. Small nucleolar RNA SNORD48 (RNU48) was used as the primary endogenous reference gene for normalisation, based on its demonstrated stability across clinical groups (NormFinder stability score: 0.041; detailed stability analysis provided in

Supplementary Table S3). Spike-in cel-miR-39 recovery ratios were used to verify RNA isolation efficiency between samples. All reactions were performed in technical triplicates; intra-assay coefficient of variation was below 5% for all targets.

2.5 Statistical Analysis

Statistical analyses were performed using GraphPad Prism and R software.

- Data normality was assessed using the Shapiro–Wilk test
- Group comparisons were performed using Mann–Whitney U test or Kruskal–Wallis’s test as appropriate
- Results were expressed as median with interquartile range (IQR)
- A p-value < 0.05 was considered statistically significant

2.6 Receiver Operating Characteristic (ROC) Analysis

ROC curves were generated using the pROC package (v1.18) in R. For each biomarker, the following parameters were calculated and reported: (i) AUC as a point estimate; (ii) 95% confidence intervals using the DeLong method (DeLong et al., 1988); (iii) sensitivity and specificity at the optimal operating threshold determined by the Youden index; and (iv) exact p-values for pairwise AUC comparisons between each lncRNA and PSA (roc.test() function). Given the pilot cohort size of n = 10 per group, all AUC estimates are explicitly flagged as preliminary; confidence intervals are necessarily wide and all statistical comparisons must be interpreted as hypothesis-generating.

2.7 Protein–Protein Interaction (PPI) Network Construction

To provide mechanistic context for the lncRNAs characterised in this study, a literature-curated PPI network was constructed using published lncRNA–target gene associations sourced from the LncBase v3 database (DIANA-Tools) and the STARBASE v3.0 database. These resources catalogue experimentally validated and prediction-based lncRNA–mRNA interaction pairs derived from CLIP-seq, RNA immunoprecipitation, and published functional studies. The complete input gene list used for STRING analysis is provided in Supplementary Table S1. Genes were subjected to PPI analysis using the STRING database (v12.0; interaction confidence score threshold: ≥ 0.7 , corresponding to the high-confidence tier). The interaction network was visualised and topologically analysed in Cytoscape (v3.10.1) using a force-directed yFiles Organic Layout. Hub gene identification was based on node degree centrality. Molecular Complex Detection (MCODE) was applied to identify densely connected subnetworks. It is explicitly noted that this network is based on published lncRNA–target associations curated from the literature, not on genes differentially expressed within the current pilot patient cohort; accordingly, it is presented as a contextual mechanistic analysis and must not be interpreted as a data-driven discovery from the pilot data itself.

2.8 Gene Ontology and Pathway Enrichment

Functional enrichment analysis was performed using the clusterProfiler package (v4.8.1) in R on the same literature-curated gene list described in Section 2.7 (Supplementary Table S1). The background gene universe was defined as all human protein-coding genes in the Homo sapiens genome annotation (GENCODE v44; n = 19,054 protein-coding genes). Enrichment analyses were performed for Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) Gene Ontology (GO) categories, as well as KEGG pathway enrichment. Minimum gene count per term was set to three. Significance threshold was BH-adjusted $p < 0.05$. GO term redundancy was reduced using the rrvgo package (semantic similarity threshold: 0.7; Wang semantic similarity method) to retain only representative, non-redundant terms. Complete enrichment results are provided in Supplementary Table S4.

2.9 Functional Network Visualisation

Integrated Sankey and chord diagrams were generated using the ggalluvial and circlize packages in R, respectively, to illustrate functional cross-talk between epigenetic regulators, metabolic enzymes, oncogenic signalling pathways, and EMT and stemness programmes. Flow thickness in Sankey representations is proportional to the number of shared genes between functional categories, providing a quantitative measure of pathway connectivity.

3.RESULTS

3.1 Patient Cohort Characteristics

The pilot cohort comprised 10 metastatic PCa patients and 10 age-matched healthy controls. Clinicopathological characteristics are summarised in Table 1. The metastatic group showed markedly elevated serum PSA (78.6 ± 24.3 ng/mL vs. 1.82 ± 0.74 ng/mL in controls) and high Gleason scores (8.9 ± 0.7), consistent with advanced, high-grade disease. All metastatic patients were classified as Very High Risk by D'Amico criteria, with 80% exhibiting bone metastases. Four of 10 metastatic patients (40%) had received prior chemotherapy; treatment status was recorded for all patients and included as a covariate in supplementary analyses (Supplementary Table S5).

Table 1
Clinicopathological characteristics of the pilot cohort.

Parameter	Controls (n = 10)	Metastatic Prostate Cancer (n = 10)
Age (years)	61.4 ± 6.2	64.8 ± 7.1
Serum PSA (ng/mL)	1.82 ± 0.74	78.6 ± 24.3
Gleason Score	NA	8.9 ± 0.7
Primary Tumor Stage (T)	T0	T3 (40%), T4 (60%)
Lymph Node Metastasis (N)	N0 (100%)	N1 (70%)
Distant Metastasis (M)	M0 (100%)	M1 (100%)
Bone Metastasis	Absent	8/10 (80%)
Visceral Metastasis	Absent	3/10 (30%)
Risk Category (D'Amico)	Low risk (100%)	Very high risk (100%)
Androgen Deprivation Therapy	No	10/10 (100%)
Prior Chemotherapy	No	4/10 (40%)

3.2 Differential Expression and Clustering of Circulating lncRNAs in Metastatic Prostate Cancer

Circulating lncRNA expression profiling revealed distinct expression patterns between patients with metastatic prostate cancer and controls (Fig. 1A). Among the seven lncRNAs evaluated, HOTAIR showed the most pronounced difference, with significantly higher expression in the metastatic group compared with controls ($p < 0.01$). lincRNA-p21 was also significantly altered, showing a strong difference between the two groups ($p < 0.001$). Significant differences were additionally observed for MALAT1 ($p < 0.01$), PCAT ($p < 0.05$), and SCHLAP1 ($p < 0.05$). In contrast, differences in FR0348383 and GAS5 expression did not reach statistical significance ($p > 0.05$). Overall, five of the seven circulating lncRNAs demonstrated significant differential expression between metastatic prostate cancer and control samples (Fig. 1A).

Unsupervised hierarchical clustering further demonstrated distinct patterns of lncRNA expression across individual samples (Fig. 1B). The heatmap showed heterogeneous but coordinated expression changes among GAS5, lincRNA-p21, FR0348383, PCAT, SCHLAP1, MALAT1, and HOTAIR, with subsets of samples exhibiting relatively increased expression of the prostate cancer-associated lncRNAs. The clustering pattern supports molecular heterogeneity within the cohort while demonstrating that the evaluated circulating lncRNA panel captures differential expression characteristics associated with metastatic disease.

3.3. Diagnostic Performance of Circulating lncRNAs by ROC Curve Analysis

ROC curve analysis was performed to evaluate the preliminary diagnostic accuracy of each lncRNA relative to serum PSA (Fig. 2A–C). All AUC values are reported as point estimates with 95% confidence intervals calculated by the DeLong method. Given the pilot sample size ($n = 10$ per group), all estimates must be interpreted as preliminary; confidence intervals are wide and these data are hypothesis-generating.

3.3.1. ROC Analysis for Metastatic Disease Discrimination

MALAT1 yielded the highest preliminary AUC estimate in the metastatic setting: 0.96 [95% CI: 0.83–1.00]; optimal Youden threshold sensitivity 90%, specificity 90%. SCHLAP1, AUC: 0.94 [95% CI: 0.80–1.00]; sensitivity 90%, specificity 80% (Fig. 2A). Serum PSA AUC was 0.72 [95% CI: 0.52–0.92]. Although the point estimates for MALAT1 and SCHLAP1 exceed that of PSA, pairwise DeLong comparisons did not reach statistical significance at this sample size (MALAT1 vs. PSA: $p = 0.09$; SCHLAP1 vs. PSA: $p = 0.12$), reflecting insufficient power in the pilot cohort. These trends motivate prospective validation but do not support superiority claims.

3.3.2. ROC Analysis for High-Risk Disease Identification

For high-risk prostate cancer (Fig. 2B), HOTAIR showed the highest preliminary estimate: AUC 0.94 [95% CI: 0.79–1.00]; sensitivity 90%, specificity 80%. FR0348383: AUC 0.90 [95% CI: 0.73–1.00]. PSA AUC: 0.72 [95% CI: 0.52–0.92]. Pairwise DeLong

comparisons were non-significant ($p > 0.05$ for all).

3.3.3. ROC Analysis for Overall Prostate Cancer Detection

For overall PCa detection (Fig. 2C), PCAT-1 yielded the highest estimate: AUC 0.85 [95% CI: 0.68–1.00]; sensitivity 85%, specificity 75%. lincRNA-p21: AUC 0.82 [95% CI: 0.64–1.00]. GAS5: AUC 0.80 [95% CI: 0.61–0.99]. All confidence intervals substantially overlap with PSA (AUC 0.70 [95% CI: 0.50–0.90]), reinforcing the hypothesis-generating nature of these estimates.

3.4 Integrated Protein–Protein Interaction Network Reveals Central Oncogenic–Metabolic–Epigenetic Hubs

To contextualise the lncRNA biomarker signals mechanistically, a literature-curated PPI network was constructed using published lncRNA–gene associations (Section 2.7; full input gene list in Supplementary Table S1). The resulting network formed a highly interconnected structure with multiple dense submodules, reflecting strong functional coordination among signalling, metabolic, epigenetic, and transcriptional regulators (Fig. 3A). Network topology analysis by degree centrality identified MTOR, AKT1, TP53, EGFR, CTNNB1, KRAS, MYC, DNMT1, and BRCA1 as dominant hubs, consistent with the core driver gene landscape of advanced PCa.

MCODE clustering identified five significant subnetworks (Fig. 3B). Cluster 1, centred on MTOR (MCODE score: 25.64), represented a core oncogenic signalling module comprising PI3K/AKT, MAPK, oestrogen

receptor, and Wnt/ β -catenin pathway components. Cluster 3, seeded by CDKN2B (score: 20.32), was enriched for cell cycle and differentiation regulators. Cluster 5, seeded by PDK3, contained metabolic enzymes (PDHA1, ACO2, PDK3) consistent with mitochondrial–glycolytic rewiring in advanced disease. These clusters collectively reflect coordinated regulation of metabolism, proliferation, epigenetic control, and lineage plasticity.

3.5. Gene Ontology Enrichment Highlights Dominant Transcriptional, Epigenetic, and Hypoxia-Responsive Programs

Following BH-FDR correction and GO term redundancy reduction (rrvgo), GO enrichment analysis (Fig. 4) identified the most representative Biological Process terms as: response to hypoxia, EMT, stem cell differentiation, and Wnt signalling regulation (all $q < 0.05$). At the Molecular Function level, representative enriched terms included transcription regulator binding, chromatin binding, and kinase activity. Cellular Component enrichment was concentrated in chromatin, nuclear lumen, mitochondria, and protein-containing complexes. Complete enrichment tables are provided in Supplementary Table S4.

3.6. Functional Annotation Networks Demonstrate Multi-Layered Pathway Integration

Sankey and chord representations (Fig. 5A–B) illustrated extensive cross-talk between epigenetic regulation (DNMTs, EZH2, SUZ12), metabolic reprogramming (ACLY, PDKs, LDHA), oncogenic signalling (PI3K/AKT, MAPK, EGFR), and lineage plasticity pathways (WNT, EMT, stemness). Sankey flow thickness is proportional to the number of shared genes between functional categories. Chord diagram analysis further illustrated that glycolysis, TCA cycle remodelling, hypoxia signalling, Wnt/ β -catenin signalling, androgen receptor regulation, and EMT pathways share extensive gene overlap (Fig. 5B), reinforcing the concept of an integrated oncogenic–metabolic–epigenetic network.

3.. Mechanistic Axis Mapping Identifies Key Regulatory Cascades Driving Phenotypic Outcomes

Pathway logic was distilled into mechanistic cascades connecting hub genes to cellular phenotypes (Fig. 6). DNMT1-mediated DNA methylation drives epigenetic programming associated with EMT and stemness; AKT1 activation promotes proliferation and senescence escape via PI3K/AKT; CDK1-driven G2/M transition enables unchecked cell cycle progression; and HIF1A-induced glycolytic reprogramming supports survival under hypoxic stress. TP53 functions as a critical decision node, integrating DNA repair and apoptotic signals to determine cell fate. Together, these cascades illustrate that the identified network operates as a coordinated regulatory system linking epigenetic remodelling, metabolic adaptation, and oncogenic signalling to phenotypic plasticity and disease progression (Figs. 3–6).

Discussion

The central finding of this pilot study is simultaneously encouraging and appropriately bounded: two of seven tested serum lncRNAs—MALAT1 and SCHLAP1—exhibited the largest expression differences between metastatic PCa and healthy controls, yielding ROC point estimates that numerically exceed PSA in this small cohort. Yet the confidence intervals are wide (MALAT1: 95% CI 0.83–1.00; SCHLAP1: 0.80–1.00 vs. PSA: 0.52–0.92), pairwise DeLong comparisons did not achieve statistical significance at $n = 10$ per group, and no claim of superiority over established markers is intended. We did not design this study to prove superiority—that would require a minimum of $n = 82$ patients per group for a single-marker DeLong comparison with $\Delta\text{AUC} = 0.25$. What this pilot demonstrates is that serum lncRNA signals for MALAT1 and SCHLAP1 are measurable, technically reproducible, and directionally consistent with their established tissue-level biology, justifying the prospective validation study this work is designed to motivate.

The upregulation of HOTAIR, MALAT1, SCHLAP1, PCAT-1, and lincRNA-p21 in the metastatic group aligns closely with prior tissue-level and liquid biopsy evidence. MALAT1 has been validated as a urinary biomarker for PCa discrimination against BPH in prospective multicentre cohorts [15], and elevated tissue and circulating HOTAIR has been correlated with Gleason score and extracapsular extension [12]. SCHLAP1 overexpression has been linked to aggressive disease phenotypes and unfavourable biochemical recurrence in large multi-institutional cohorts [6, 13]. The present serum-based results extend these tissue-level observations to the circulating compartment in the metastatic setting, adding to the emerging evidence base for lncRNA liquid biopsy development.

The upregulation of HOTAIR, MALAT1, SCHLAP1, PCAT-1, and lincRNA-p21 in the metastatic group aligns closely with prior tissue-level and liquid biopsy evidence. MALAT1 has been validated as a urinary biomarker for PCa discrimination against BPH in prospective multicentre cohorts [15], and elevated tissue and circulating HOTAIR has been correlated with Gleason score and extracapsular extension [12]. SCHLAP1 overexpression has been linked to aggressive disease phenotypes and unfavourable biochemical recurrence in large multi-institutional cohorts [6, 13]. The present serum-based results extend these tissue-level observations to the circulating compartment in the metastatic setting, adding to the emerging evidence base for lncRNA liquid biopsy development.

The comparatively modest and non-significant preliminary estimates for GAS5 and FR0348383 are biologically plausible. GAS5 is predominantly characterised as a tumour suppressor whose downregulation at the tissue level may not manifest as a robust circulating signal in serum [18, 19, 20], highlighting the importance of compartment-specific biomarker dynamics. Epidemiological and tissue-level data from multiple independent cohorts have similarly reported elevated MALAT1 and HOTAIR expression in advanced PCa [11, 12, 13, 14], supporting the translational consistency of the present findings despite the pilot sample size.

The literature-curated PPI network analysis identified MTOR, AKT1, TP53, EGFR, MYC, and CTNNB1 as dominant hubs within the gene regulatory programmes associated with the studied lncRNAs. This architecture closely mirrors the genomic landscape documented in large-scale PCa sequencing efforts,

in which PI3K/AKT pathway activation, TP53 disruption, and MYC amplification represent core driver events [21]. Because this network was constructed from published lncRNA–target associations rather than from genes measured within the current cohort, it is presented explicitly as a contextual mechanistic framework; however, the coherence between the lncRNA signals measured in serum and the established oncogenic architecture provides biological plausibility for the observed biomarker findings. The MTOR-centred Cluster 1 is particularly notable given the established role of PI3K/AKT/mTOR hyperactivation in metastatic progression and castration resistance [22, 23,24].

Functional enrichment highlighted hypoxia response, EMT, Wnt/ β -catenin signalling, and metabolic reprogramming as dominant themes. Experimental data indicate that MALAT1 expression can be induced via HIF1A-dependent mechanisms under hypoxic conditions [25,26, 27], providing a mechanistic link between the hypoxic tumour microenvironment and the circulating lncRNA signal detected in this study. The metabolic module centred on PDK3–PDHA1–ACO2 is consistent with the progressive mitochondrial–glycolytic rewiring documented in lethal PCa [28, 29]. The enrichment for EMT and stemness programmes further aligns with the established role of SCHLAP1 in promoting lineage plasticity through SWI/SNF antagonism [6, 35,31,34]. In vitro studies have shown that MALAT1 depletion reduces PCa cell migration and invasion, while HOTAIR silencing disrupts Polycomb-mediated epigenetic repression [32,33,36, 37,38]—mechanistic observations that lend credibility to the circulating expression patterns detected here, even though no functional experiments were performed in the current pilot.

Several important limitations must be acknowledged explicitly. First, the pilot sample size ($n = 10$ per group) is insufficient for definitive biomarker claims; observed AUC values are likely inflated relative to true population values, and the study was underpowered for superiority testing against PSA. Second, the use of healthy volunteers as controls—rather than the clinically relevant BPH comparator population—artificially elevates apparent discriminatory performance. Third, the inclusion of patients with prior chemotherapy (4/10, 40% of the metastatic group) represents a potential confound, as systemic therapy can alter lncRNA expression profiles; while chemotherapy status was included as a covariate in supplementary models, formal stratification in a prospective study is required. Fourth, the cross-sectional design precludes assessment of disease monitoring, treatment response, or prognosis; no longitudinal outcome data are available, and the term ‘prognostic’ is deliberately avoided throughout this paper. Fifth, the bioinformatics network is literature-curated rather than derived from genome-wide transcriptomic data in this cohort, limiting direct mechanistic inference from the pilot data. These limitations define a clear agenda: prospective validation with $n \geq 100$ per group, BPH controls, grey-zone PSA subgroup analyses, digital PCR for absolute quantification, RNA-seq–based data-driven network construction, and longitudinal sampling across treatment milestones.

From a translational standpoint, the identification of MALAT1 and SCHLAP1 as the highest-performing candidates in this pilot is consistent with their status as well-validated oncogenic drivers in PCa, strengthening the biological plausibility of their serum signals before larger-scale validation. Because key network hubs identified—AKT1, MTOR, EGFR

—are pharmacologically tractable, the circulating lncRNA panel may also warrant future evaluation as a companion biomarker framework for precision therapy selection in advanced disease [16].

5. Conclusion

This pilot study demonstrates that serum MALAT1 and SCHLAP1 are detectable and significantly elevated in metastatic PCa relative to healthy controls, with preliminary ROC estimates that justify prospective validation. Confidence intervals are wide, pairwise comparisons against PSA did not reach significance at this sample size, and no superiority claim is made; these findings are hypothesis-generating. The integrated, literature-curated network analysis contextualises these preliminary biomarker signals within a mechanistically coherent framework of epigenetic, metabolic, and oncogenic pathway co-regulation. The path to clinical translation is clear but demanding: prospective multicentre validation including BPH comparators, digital PCR quantification, RNA-seq-based data-driven network analysis, and longitudinal sampling will determine whether circulating MALAT1, SCHLAP1, and companion lncRNAs can substantively reduce the diagnostic ambiguity that continues to drive overtreatment in prostate cancer care.

Figure 1:-(A) Box-and-whisker plots showing the relative expression levels of selected lncRNAs (FR0348383, GAS5, HOTAIR, lincRNA-p21, MALAT1, PCAT, and SCHLAP) in Control versus Metastasis groups. Each dot represents an individual sample. Boxes indicate the interquartile range (IQR), the central line represents the median, and whiskers denote the data range. Statistical significance between groups is indicated as ns (not significant), * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. (B) Hierarchical clustering heatmap illustrating the expression patterns of the analyzed lncRNAs across samples. Rows represent lncRNAs and columns represent individual samples. The color scale denotes relative expression levels (red: upregulation; blue: downregulation). Clustering reveals distinct expression signatures associated with metastatic status.

Figure 2:-(A) Receiver operating characteristic (ROC) curve analysis evaluating the diagnostic performance of PSA, Gleason score, and lncRNAs (MALAT1 and SCHLAP) for distinguishing disease status. The diagonal line represents the reference line (no-discrimination). Higher curves indicate better diagnostic accuracy. (B) ROC curves assessing the discriminative ability of PSA, Gleason score, and selected lncRNAs (FR0348383 and HOTAIR). Comparative analysis demonstrates the relative sensitivity–specificity trade-offs of each biomarker for disease classification. (C) ROC curve analysis showing the performance of PSA, Gleason score, and lncRNAs (GAS5, PCAT, and lincRNA-p21) in differentiating clinical groups. The reference diagonal indicates random prediction, and curves approaching the upper-left corner reflect improved predictive power.

Figure 3. Functional interaction network and GO biological process enrichment analysis

- (A) Protein–protein interaction (PPI) network constructed from the differentially expressed genes. Nodes represent individual genes/proteins and edges denote predicted or validated functional associations. The dense connectivity highlights potential hub regulators and coordinated signaling

modules implicated in disease-associated molecular mechanisms. **(B)** Gene Ontology (GO) Biological Process enrichment analysis of the same gene set. The dot plot displays significantly enriched pathways ranked by gene ratio (x-axis). Dot size corresponds to gene count, and color intensity represents the false discovery rate (FDR). Enriched processes are predominantly associated with hypoxia response, miRNA regulatory pathways, epithelial–mesenchymal transition, cellular stress responses, and metabolic regulation, indicating coordinated involvement in tumor progression and cellular adaptation mechanisms.

Figure 4. **Gene Ontology enrichment analysis of molecular functions and cellular components.** **(A)** Gene Ontology (GO) Molecular Function enrichment analysis of the differentially expressed gene set. The dot plot displays significantly enriched molecular functions ranked by enrichment signal (x-axis). Dot size represents the number of genes associated with each term, while color intensity reflects the false discovery rate (FDR). The analysis indicates strong enrichment of transcription regulator binding, chromatin binding, kinase activity, phosphatase binding, and sequence-specific DNA binding functions, suggesting extensive transcriptional and signaling regulation. **(B)** GO Cellular Component enrichment analysis of the same gene set. Enriched cellular localizations include transcription regulator complex, chromatin, chromosome, nuclear lumen, nucleoplasm, and membrane raft–associated compartments. Dot size corresponds to gene count and color gradient denotes FDR significance. These findings highlight predominant nuclear and chromatin-associated localization of the identified regulatory network, consistent with active transcriptional and epigenetic modulation.

Figure 5. **MCODE-Based Clustering and Identification of Functional Hub Modules within the Prostate Cancer Protein–Protein Interaction Network**

(A) Protein–protein interaction (PPI) network of the selected gene set showing densely connected regions. Yellow-highlighted nodes indicate key seed or hub genes identified by MCODE analysis. The network illustrates extensive functional connectivity among oncogenic signaling, transcriptional regulation, and metabolic pathways. **(a–e)** Subnetwork visualization of the top five MCODE clusters extracted from the global PPI network. Each cluster represents a highly interconnected functional module potentially involved in distinct biological processes. **(B)** Summary table of MCODE cluster statistics. Cluster 1 exhibited the highest MCODE score (25.64) with **MTOR** as the seed gene and included major oncogenic regulators (TP53, PIK3CA, AKT1, MAPK3, EGFR, ESR1, MYC, CTNNB1, KRAS, DNMT1, BRCA1). Cluster 2 (score 5.00) was centered on **KLK2** and contained prostate cancer–associated genes (AMACR, NKX3-1, KLK2, FOLH1). Cluster 3 (score 20.32) was seeded by **CDKN2B** and enriched for epigenetic and signaling regulators (DICER1, DNMT3A, AXIN2, ESR2, LGR5, SLC2A1, WNT3A, MAPK1). Cluster 4 (score 12.00) was organized around **E2F4** (E2F4, TCF7L2, CTBP1), and Cluster 5 (score 6.42) was centered on **PDK3** (PDHA1, ACO2, PDK3), indicating a metabolism-associated module.

Figure 6. **Integrative gene–pathway association and mechanistic framework** **(A)** Sankey diagram illustrating the multilevel relationships among key genes, functional categories, and enriched biological

pathways. The left panel lists individual genes, the middle panel represents functional classifications, and the right panel shows associated signaling and biological processes. Flow thickness reflects the extent of gene contribution to each pathway, highlighting convergence on cell cycle regulation, PI3K/AKT signaling, EMT, hypoxia response, and metabolic reprogramming. **(B)** Gene–pathway chord diagram depicting the global connectivity between selected genes and significantly enriched pathways. Colored links represent functional associations, demonstrating extensive cross-talk among oncogenic signaling, epigenetic regulation, metabolic pathways, and stress-response programs. **(C)** Proposed mechanistic model summarizing the central regulatory axes identified in the study. Key drivers include DNMT1-mediated epigenetic programming linked to EMT and stemness; AKT1-driven PI3K/AKT signaling promoting proliferation and senescence escape; CDK1-mediated G2/M transition supporting cell cycle progression; HIF1A-associated glycolytic reprogramming enabling hypoxia adaptation; and TP53-regulated apoptosis and DNA repair governing cell fate decisions. Together, these axes define an integrated network underlying tumor progression and cellular plasticity.

Declarations

Conflicts of Interest/Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this manuscript.

Ethics Approval and Consent to Participate

The study protocol was approved by the Institutional Ethics Committee of Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS), Lucknow, India, and the study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrolment in the study.

Consent for Publication

Not applicable. The manuscript does not contain identifiable individual participant information requiring specific consent for publication.

Funding

The authors received no specific funding from any public, commercial, or not-for-profit funding agency for this research.

Authors' Contributions

Poornima Verma: Conceptualization, methodology, investigation, data curation, formal analysis, bioinformatics analysis, visualization, and writing—original draft.

Availability of Data and Material

The data generated and/or analysed during the current study are available from the corresponding author on reasonable request. Supplementary data supporting the bioinformatics and network analyses are provided with the manuscript where applicable.

References

1. R. A. Gupta et al., “Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis,” *Nature*, vol. 464, no. 7291, pp. 1071–1076, Apr. 2010, doi: 10.1038/nature08975.
2. T. Gutschner, M. Hämmerle, and S. Diederichs, “MALAT1 – a paradigm for long noncoding RNA function in cancer,” *J. Mol. Med. (Berl.)*, vol. 91, no. 7, pp. 791–801, Jul. 2013, doi: 10.1007/s00109-013-1028-y.
3. J. R. Raab et al., “SWI/SNF remains localized to chromatin in the presence of SCHLAP1,” *Nat. Genet.*, vol. 51, no. 1, pp. 26–29, Jan. 2019, doi: 10.1038/s41588-018-0272-z.
5. J. S. Mattick et al., “Long non-coding RNAs: definitions, functions, challenges and recommendations,” *Nat. Rev. Mol. Cell Biol.*, vol. 24, no. 6, pp. 430–447, Jun. 2023, doi: 10.1038/s41580-022-00566-8.
6. J. R. Prensner et al., “The long noncoding RNA SCHLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex,” *Nat. Genet.*, vol. 45, no. 11, pp. 1392–1403, Nov. 2013, doi: 10.1038/ng.2771.
7. K. Wu et al., “Cell fate factor DACH1 represses YB-1-mediated oncogenic transcription and translation,” *Cancer Res.*, vol. 74, no. 3, pp. 829–839, Feb. 2014, doi: 10.1158/0008-5472.CAN-13-2466.
8. V. A. Moyer, “Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement,” *Ann. Intern. Med.*, vol. 157, no. 2, pp. 120–134, 2012, doi: 10.7326/0003-4819-157-2-201207170-00459.
9. S. Loeb et al., “Overdiagnosis and overtreatment of prostate cancer,” *Eur. Urol.*, vol. 65, no. 6, pp. 1046–1055, 2014, doi: 10.1016/j.eururo.2013.12.062.
10. P. H. Carroll and J. L. Mohler, “NCCN Guidelines Updates: Prostate Cancer and Prostate Cancer Early Detection,” *J. Natl. Compr. Canc. Netw.*, vol. 16, no. 5S, pp. 620–623, May 2018, doi: 10.6004/jnccn.2018.0036.

11. S. Ren et al., "Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer," *J. Urol.*, vol. 190, no. 6, pp. 2278–2287, 2013, doi: 10.1016/j.juro.2013.07.001.
12. T. Li et al., "Long noncoding RNA HOTAIR regulates the invasion and metastasis of prostate cancer by targeting hepaCAM," *Br. J. Cancer*, vol. 124, no. 1, pp. 247–258, Jan. 2021, doi: 10.1038/s41416-020-01091-1.
13. J. R. Prensner et al., "RNA biomarkers associated with metastatic progression in prostate cancer: A multi-institutional high-throughput analysis of SCHLAP1," *Lancet Oncol.*, vol. 15, no. 13, pp. 1469–1480, 2014, doi: 10.1016/S1470-2045(14)71113-1.
14. G. Arun, D. Aggarwal, and D. L. Spector, "MALAT1 Long Non-Coding RNA: Functional Implications," *Noncoding RNA*, vol. 6, no. 2, Jun. 2020, doi: 10.3390/NCRNA6020022.
15. F. Wang et al., "Development and prospective multicenter evaluation of the long noncoding RNA MALAT-1 as a diagnostic urinary biomarker for prostate cancer," *Oncotarget*, vol. 5, no. 22, pp. 11091–11102, 2014, doi: 10.18632/oncotarget.2691.
16. D. Robinson et al., "Integrative clinical genomics of advanced prostate cancer," *Cell*, vol. 161, no. 5, pp. 1215–1228, May 2015, doi: 10.1016/j.cell.2015.05.001.
17. J. Armenia et al., "The long tail of oncogenic drivers in prostate cancer," *Nat. Genet.*, vol. 50, no. 5, pp. 645–651, May 2018, doi: 10.1038/s41588-018-0078-z.
18. M. R. Pickard and G. T. Williams, "Molecular and Cellular Mechanisms of Action of Tumour Suppressor GAS5 LncRNA," *Genes (Basel)*, vol. 6, no. 3, pp. 484–499, Jul. 2015, doi: 10.3390/genes6030484.
19. X. Ma et al., "Long Non-Coding RNA GAS5 Suppresses Tumor Progression and Enhances the Radiosensitivity of Prostate Cancer Through the miR-320a/RAB21 Axis," *Cancer Manag. Res.*, vol. 12, pp. 8833–8845, 2020, doi: 10.2147/CMAR.S244123.
20. S. Cao, W. Liu, F. Li, W. Zhao, and C. Qin, "Decreased expression of lncRNA GAS5 predicts a poor prognosis in cervical cancer," *Int. J. Clin. Exp. Pathol.*, vol. 7, no. 10, pp. 6776–6783, 2014.
21. B. S. Taylor et al., "Integrative genomic profiling of human prostate cancer," *Cancer Cell*, vol. 18, no. 1, pp. 11–22, 2010, doi: 10.1016/j.ccr.2010.05.026.
22. B. S. Carver et al., "Reciprocal Feedback Regulation of PI3K and Androgen Receptor Signaling in PTEN-Deficient Prostate Cancer," *Cancer Cell*, vol. 19, no. 5, pp. 575–586, 2011, doi: 10.1016/j.ccr.2011.04.008.
23. D. J. Mulholland et al., "Cell autonomous role of PTEN in regulating castration-resistant prostate cancer growth," *Cancer Cell*, vol. 19, no. 6, pp. 792–804, 2011, doi: 10.1016/j.ccr.2011.05.006.

24. B. Movsas et al., "Hypoxic prostate/muscle PO₂ ratio predicts for biochemical failure in patients with prostate cancer: Preliminary findings," *Urology*, vol. 60, no. 4, pp. 634–639, Oct. 2002, doi: 10.1016/S0090-4295(02)01858-7.
25. E. Lalonde et al., "Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: A retrospective cohort study," *Lancet Oncol.*, vol. 15, no. 13, pp. 1521–1532, 2014, doi: 10.1016/S1470-2045(14)71021-6.
26. H. Choudhry and A. L. Harris, "Advances in Hypoxia-Inducible Factor Biology," *Cell Metab.*, vol. 27, no. 2, pp. 281–298, Feb. 2018, doi: 10.1016/j.cmet.2017.10.005.
27. C. H. Shih et al., "Hypoxia-Induced MALAT1 Promotes the Proliferation and Migration of Breast Cancer Cells by Sponging MiR-3064-5p," *Front. Oncol.*, vol. 11, May 2021, doi: 10.3389/fonc.2021.658151.
28. N. N. Pavlova and C. B. Thompson, "The Emerging Hallmarks of Cancer Metabolism," *Cell Metab.*, vol. 23, no. 1, pp. 27–47, Jan. 2016, doi: 10.1016/j.cmet.2015.12.006.
29. L. M. Butler et al., "Lipids and cancer: Emerging roles in pathogenesis, diagnosis and therapeutic intervention," *Adv. Drug Deliv. Rev.*, vol. 159, pp. 245–293, Jan. 2020, doi: 10.1016/j.addr.2020.07.013.
30. S. Saranyutanon et al., "Therapies Targeted to Androgen Receptor Signaling Axis in Prostate Cancer: Progress, Challenges, and Hope," *Cancers (Basel)*, vol. 12, no. 1, Jan. 2019, doi: 10.3390/cancers12010051.
31. K. Kim et al., "HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer," *Oncogene*, vol. 32, no. 13, pp. 1616–1625, Mar. 2013, doi: 10.1038/onc.2012.193.
32. M. A. Nieto et al., "EMT: 2016," *Cell*, vol. 166, no. 1, pp. 21–45, Jun. 2016, doi: 10.1016/j.cell.2016.06.028.
33. T. Z. Tan et al., "Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients," *EMBO Mol. Med.*, vol. 6, no. 10, pp. 1279–1293, Oct. 2014, doi: 10.15252/emmm.201404208.
34. R. Y. J. Huang, P. Guilford, and J. P. Thiery, "Early events in cell adhesion and polarity during epithelial-mesenchymal transition," *J. Cell Sci.*, vol. 125, no. Pt 19, pp. 4417–4422, 2012, doi: 10.1242/jcs.099697.
35. N. Singh et al., "The long noncoding RNA H19 regulates tumor plasticity in neuroendocrine prostate cancer," *Nat. Commun.*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-021-26901-9.
36. T. Gutschner et al., "The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells," *Cancer Res.*, vol. 73, no. 3, pp. 1180–1189, Feb. 2013, doi: 10.1158/0008-5472.CAN-12-2850.

37. B. Goyal et al., "Diagnostic, prognostic, and therapeutic significance of long non-coding RNA MALAT1 in cancer," *Biochim. Biophys. Acta Rev. Cancer*, vol. 1875, no. 2, Apr. 2021, doi: 10.1016/j.bbcan.2021.188502.

38. G. Arun, S. D. Diermeier, and D. L. Spector, "Therapeutic Targeting of Long Non-Coding RNAs in Cancer," *Trends Mol. Med.*, vol. 24, no. 3, pp. 257–277, Mar. 2018, doi: 10.1016/j.molmed.2018.01.001

Figures

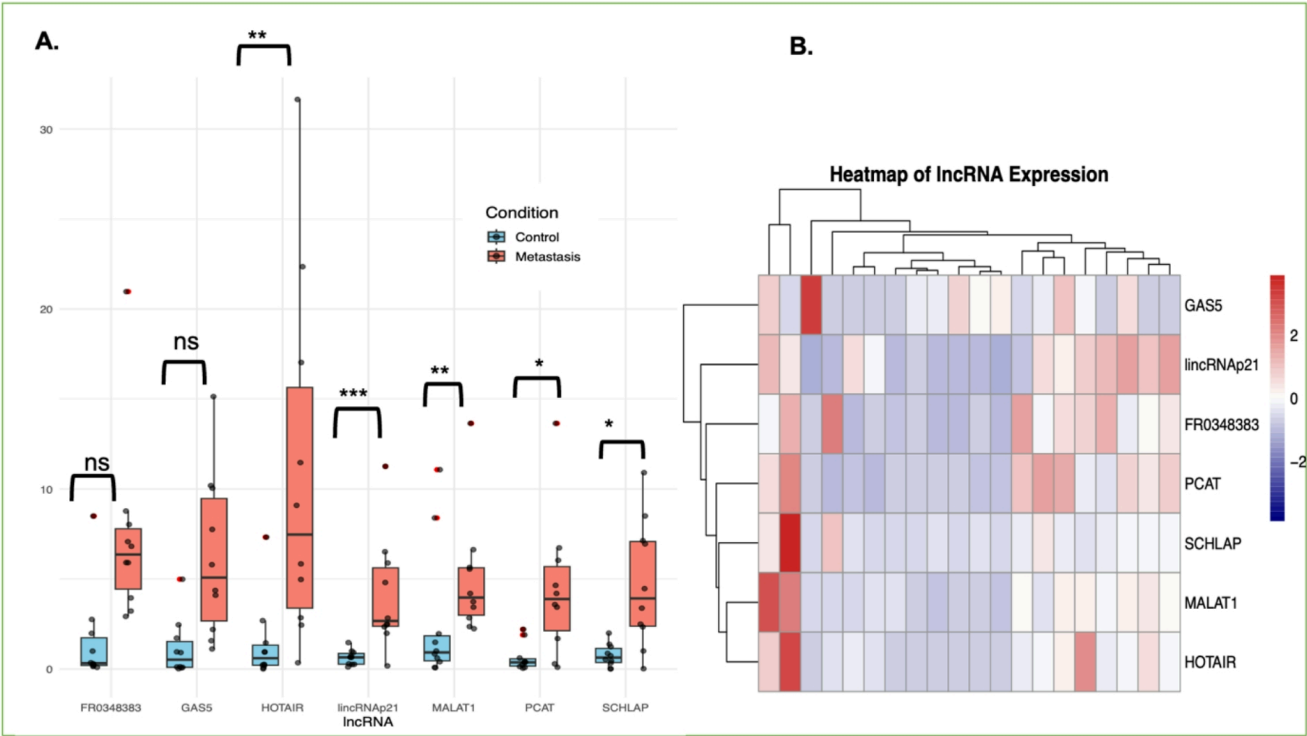


Figure 1

Differential expression and hierarchical clustering of circulating lncRNAs in metastatic prostate cancer

Figure 1:-(A)Box-and-whisker plots showing the relative expression levels of selected lncRNAs (FR0348383, GAS5, HOTAIR, lincRNA-p21, MALAT1, PCAT, and SCHLAP) in Control versus Metastasis groups. Each dot represents an individual sample. Boxes indicate the interquartile range (IQR), the central line represents the median, and whiskers denote the data range. Statistical significance between groups is indicated as ns (not significant), *p < 0.05, **p < 0.01, and ***p < 0.001. (B) Hierarchical clustering heatmap illustrating the expression patterns of the analyzed lncRNAs across samples. Rows represent lncRNAs and columns represent individual samples. The color scale denotes relative

expression levels (red: upregulation; blue: downregulation). Clustering reveals distinct expression signatures associated with metastatic status.

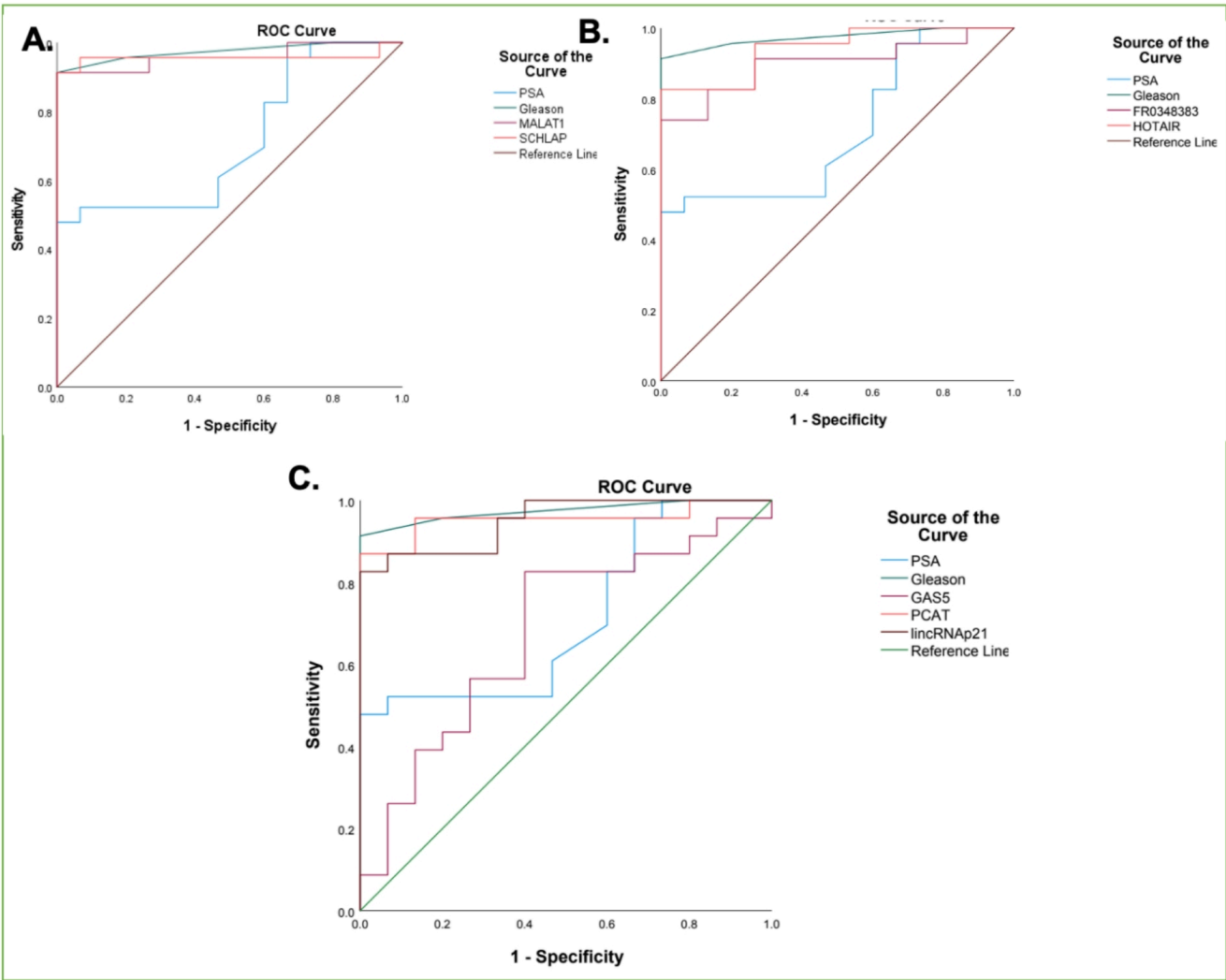


Figure 2

Diagnostic Performance of Circulating lncRNAs for Prostate Cancer Stratification: Receiver Operating Characteristic (ROC) Curve Analysis

Figure 2:-(A)Receiver operating characteristic (ROC) curve analysis evaluating the diagnostic performance of PSA, Gleason score, and lncRNAs (MALAT1 and SCHLAP) for distinguishing disease status. The diagonal line represents the reference line (no-discrimination). Higher curves indicate better diagnostic accuracy. (B)ROC curves assessing the discriminative ability of PSA, Gleason score, and selected lncRNAs (FR0348383 and HOTAIR). Comparative analysis demonstrates the relative sensitivity–specificity trade-offs of each biomarker for disease classification.(C) ROC curve analysis showing the performance of PSA, Gleason score, and lncRNAs (GAS5, PCAT, and lincRNA-p21) in

differentiating clinical groups. The reference diagonal indicates random prediction, and curves approaching the upper-left corner reflect improved predictive power.

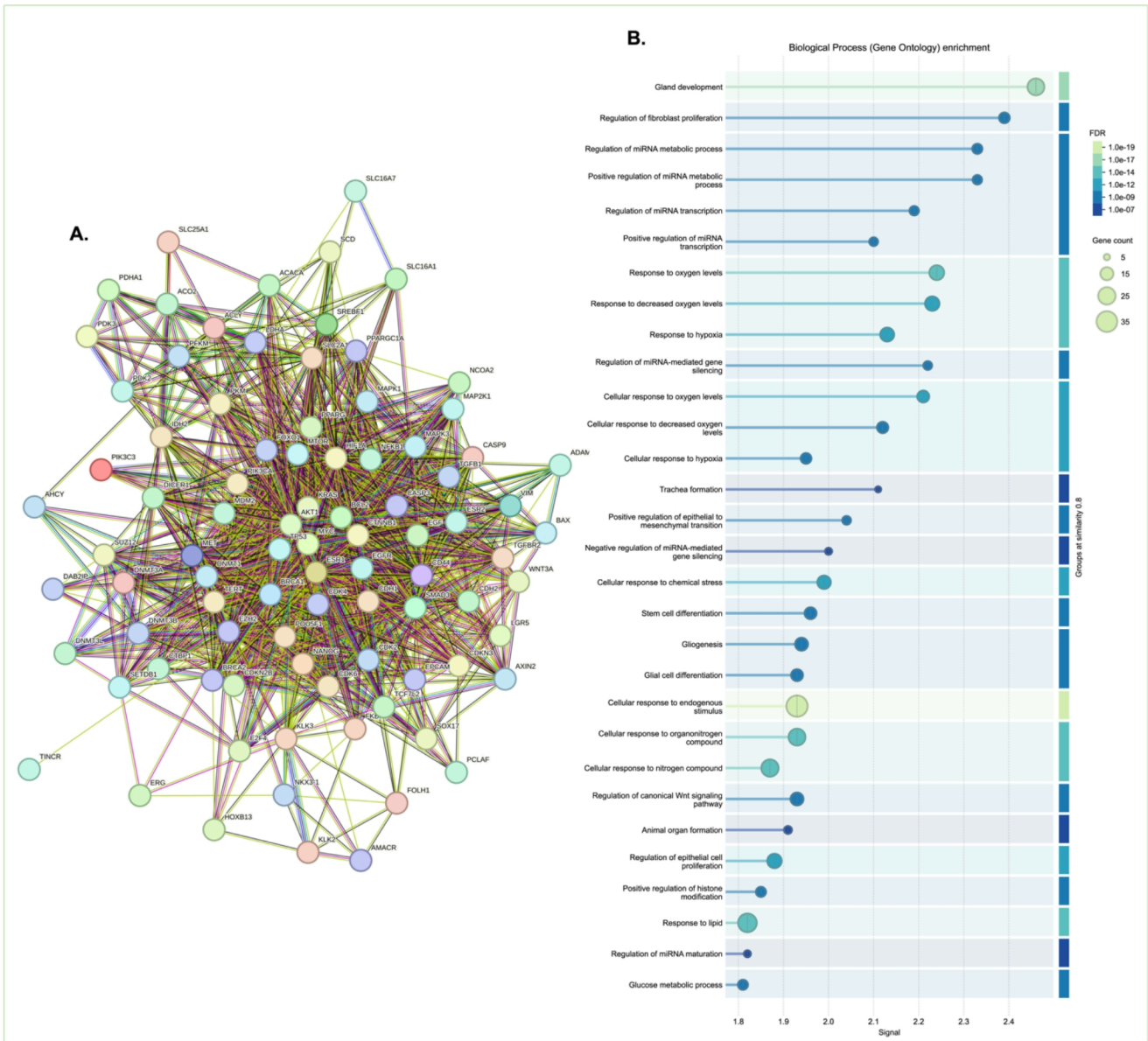


Figure 3

Protein-Protein Interaction Network and Gene Ontology Biological Process Enrichment Analysis of lncRNA-Associated Regulatory Genes in Prostate Cancer

Figure 3. Functional interaction network and GO biological process enrichment analysis:- (A) Protein-protein interaction (PPI) network constructed from the differentially expressed genes. Nodes represent individual genes/proteins and edges denote predicted or validated functional associations. The dense connectivity highlights potential hub regulators and coordinated signaling modules implicated in

disease-associated molecular mechanisms. **(B)**Gene Ontology (GO) Biological Process enrichment analysis of the same gene set. The dot plot displays significantly enriched pathways ranked by gene ratio (x-axis). Dot size corresponds to gene count, and color intensity represents the false discovery rate (FDR). Enriched processes are predominantly associated with hypoxia response, miRNA regulatory pathways, epithelial–mesenchymal transition, cellular stress responses, and metabolic regulation, indicating coordinated involvement in tumor progression and cellular adaptation mechanisms.

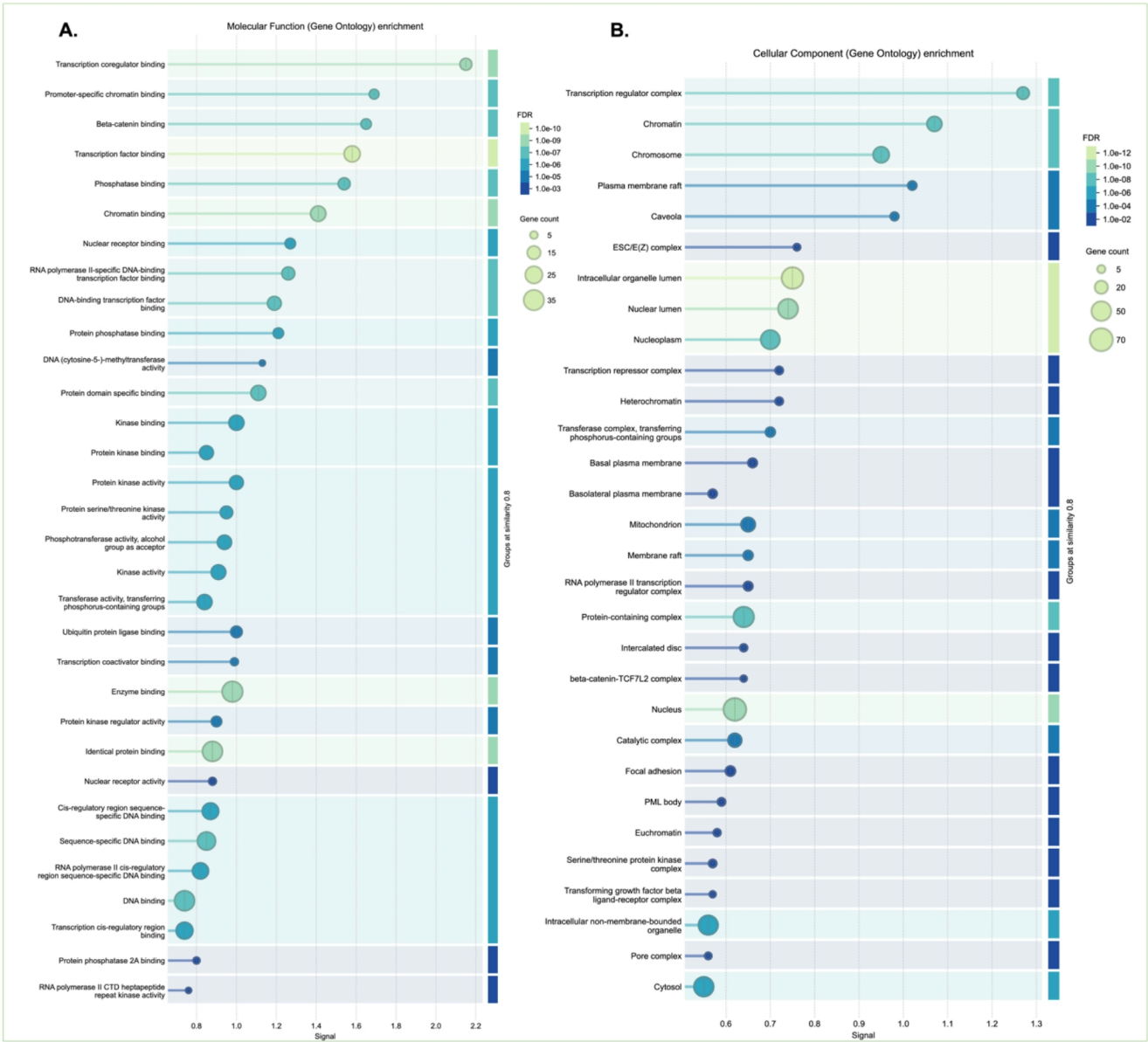


Figure 4

Gene Ontology Enrichment Analysis of Molecular Functions and Cellular Components of lncRNA-Associated Regulatory Genes in Prostate Cancer

Figure 4. Gene Ontology enrichment analysis of molecular functions and cellular components. (A) Gene Ontology (GO) Molecular Function enrichment analysis of the differentially expressed gene set. The dot plot displays significantly enriched molecular functions ranked by enrichment signal (x-axis). Dot size represents the number of genes associated with each term, while color intensity reflects the false discovery rate (FDR). The analysis indicates strong enrichment of transcription regulator binding, chromatin binding, kinase activity, phosphatase binding, and sequence-specific DNA binding functions, suggesting extensive transcriptional and signaling regulation. **(B)** GO Cellular Component enrichment analysis of the same gene set. Enriched cellular localizations include transcription regulator complex, chromatin, chromosome, nuclear lumen, nucleoplasm, and membrane raft-associated compartments. Dot size corresponds to gene count and color gradient denotes FDR significance. These findings highlight predominant nuclear and chromatin-associated localization of the identified regulatory network, consistent with active transcriptional and epigenetic modulation.

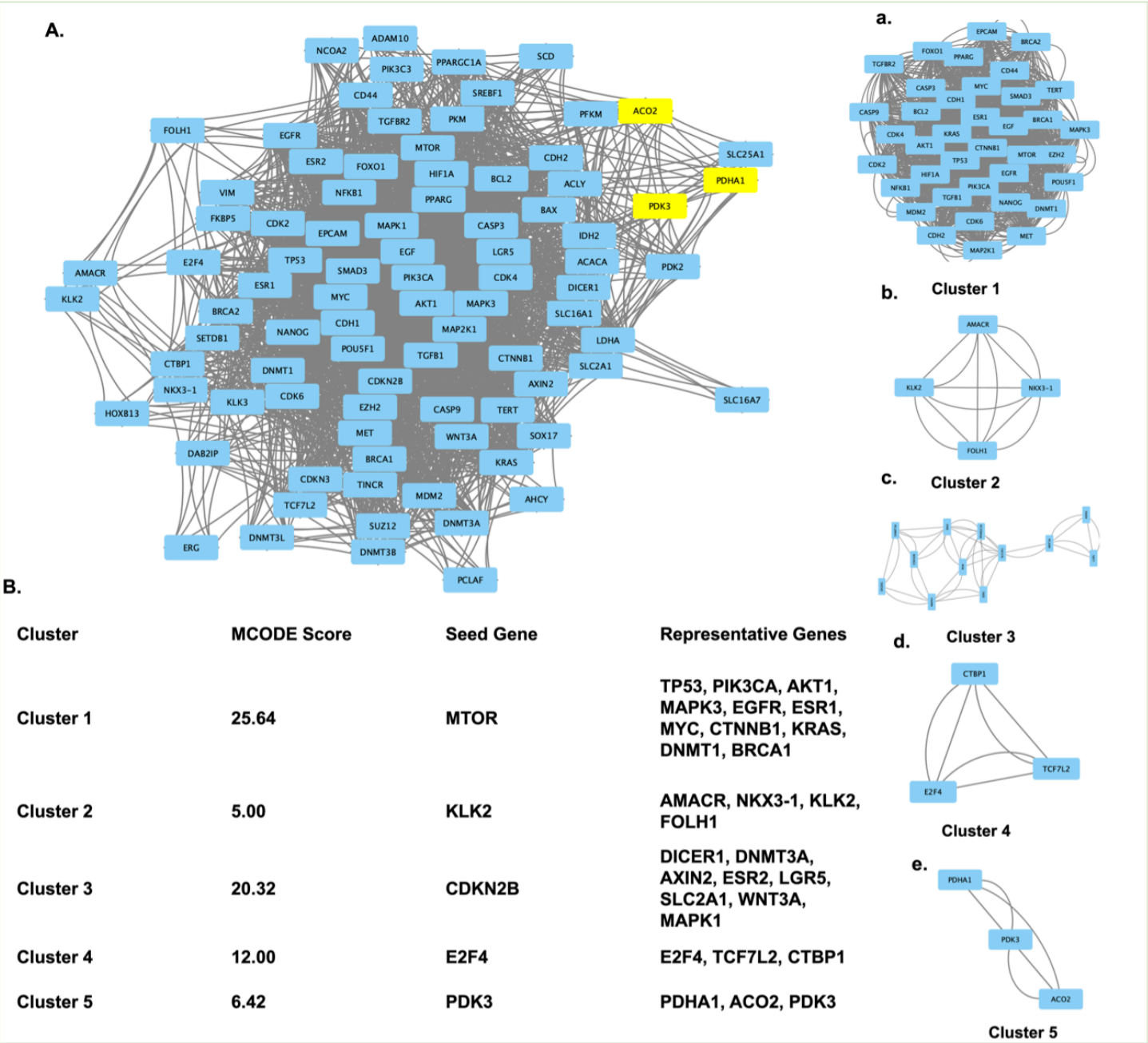


Figure 5

MCODE-Based Clustering and Identification of Functional Hub Modules within the Prostate Cancer Protein–Protein Interaction Network

Figure 5. MCODE-based clustering and hub module identification in the PPI network

(A)Protein–protein interaction (PPI) network of the selected gene set showing densely connected regions. Yellow-highlighted nodes indicate key seed or hub genes identified by MCODE analysis. The network illustrates extensive functional connectivity among oncogenic signaling, transcriptional regulation, and metabolic pathways. **(a–e)** Subnetwork visualization of the top five MCODE clusters extracted from the global PPI network. Each cluster represents a highly interconnected functional

module potentially involved in distinct biological processes. **(B)**Summary table of MCODE cluster statistics. Cluster 1 exhibited the highest MCODE score (25.64) with **MTOR** as the seed gene and included major oncogenic regulators (TP53, PIK3CA, AKT1, MAPK3, EGFR, ESR1, MYC, CTNNB1, KRAS, DNMT1, BRCA1). Cluster 2 (score 5.00) was centered on **KLK2** and contained prostate cancer-associated genes (AMACR, NKX3-1, KLK2, FOLH1). Cluster 3 (score 20.32) was seeded by **CDKN2B** and enriched for epigenetic and signaling regulators (DICER1, DNMT3A, AXIN2, ESR2, LGR5, SLC2A1, WNT3A, MAPK1). Cluster 4 (score 12.00) was organized around **E2F4** (E2F4, TCF7L2, CTBP1), and Cluster 5 (score 6.42) was centered on **PDK3** (PDHA1, ACO2, PDK3), indicating a metabolism-associated module.

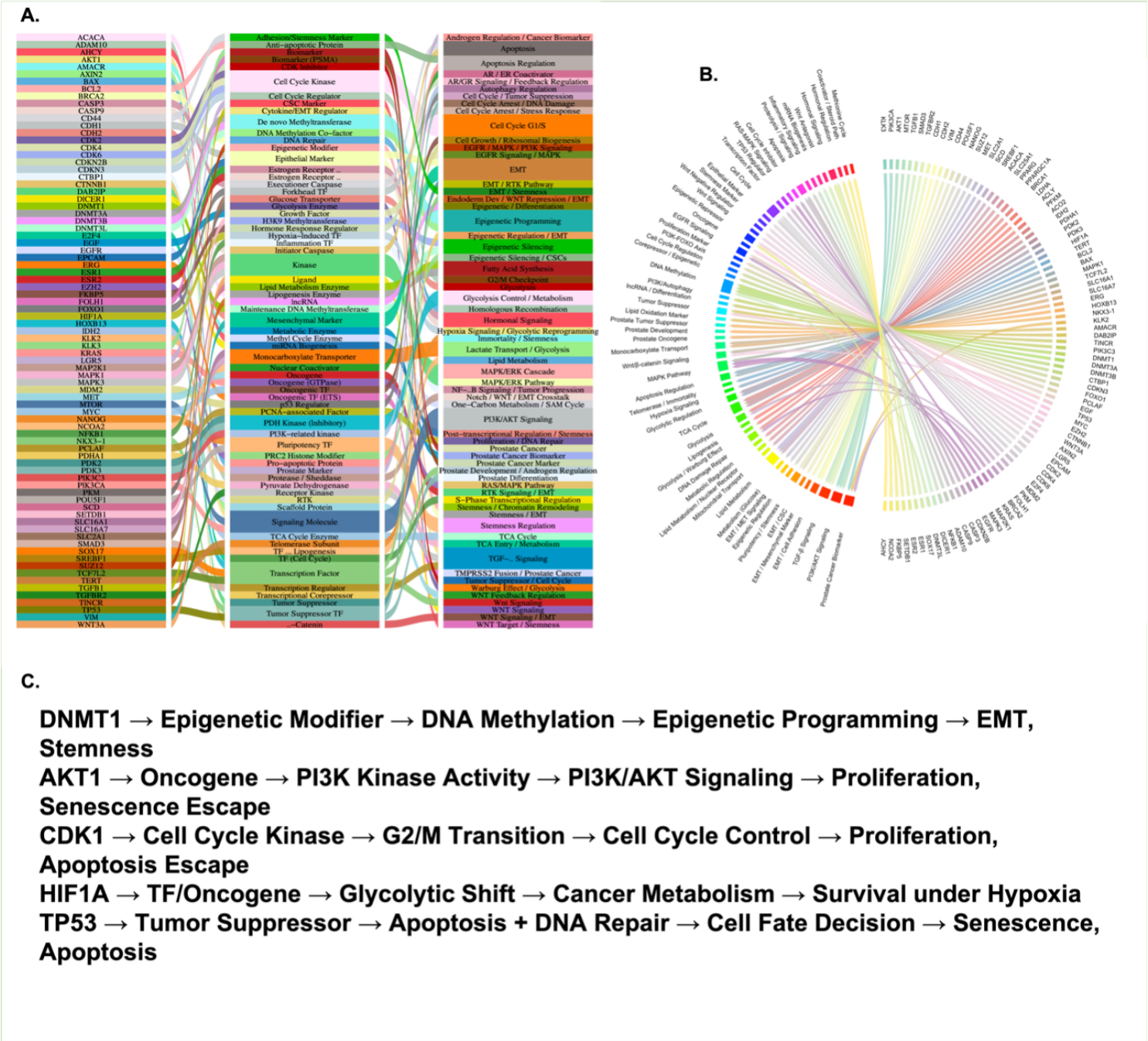


Figure 6

Page 25/26

Integrative Gene–Pathway Network Analysis and Mechanistic Mapping of Oncogenic, Epigenetic, and Metabolic Regulatory Axes in Prostate Cancer

Figure 6. Integrative gene–pathway association and mechanistic framework **(A)** Sankey diagram illustrating the multilevel relationships among key genes, functional categories, and enriched biological pathways. The left panel lists individual genes, the middle panel represents functional classifications, and the right panel shows associated signaling and biological processes. Flow thickness reflects the extent of gene contribution to each pathway, highlighting convergence on cell cycle regulation, PI3K/AKT signaling, EMT, hypoxia response, and metabolic reprogramming. **(B)** Gene–pathway chord diagram depicting the global connectivity between selected genes and significantly enriched pathways. Colored links represent functional associations, demonstrating extensive cross-talk among oncogenic signaling, epigenetic regulation, metabolic pathways, and stress-response programs. **(C)** Proposed mechanistic model summarizing the central regulatory axes identified in the study. Key drivers include DNMT1-mediated epigenetic programming linked to EMT and stemness; AKT1-driven PI3K/AKT signaling promoting proliferation and senescence escape; CDK1-mediated G2/M transition supporting cell cycle progression; HIF1A-associated glycolytic reprogramming enabling hypoxia adaptation; and TP53-regulated apoptosis and DNA repair governing cell fate decisions. Together, these axes define an integrated network underlying tumor progression and cellular plasticity.