Applied Biological Research 28(3): 324-334; (2026)
DOI: 10.48165/abr.2026.28.01.33

INTEGRATIVE IDENTIFICATION OF RECURRENT NON-CODING MUTATIONS IN MICRORNAs AND TRANSCRIPTION FACTOR BINDING SITES FOR EARLY DETECTION OF PROSTATE CANCER

Amna Amin Sethi, Faizan Siraj Qureshi, Syed Fahad Ali, Haniah Iqbal Naseem, and Tayyaba Wasi

Department of Biomedical Engineering, NED University of Engineering & Technology, LEJ Campus, Karachi – 74800, Sindh (Pakistan)
*****e-mail: *amnasethi@neduet.edu.pk*

(Received 23 March, 2026; Accepted 8 August, 2026)

ABSTRACT

Prostate cancer remains a major global health challenge, and the largely unexplored non-coding genome—including microRNAs (miRNAs) and transcription factor binding sites (TFBS) — offers critical insights into its early molecular events. This study is distinct in integration of computational analyses of both TFBS and miRNAs to identify highly recurrent mutations within non-coding regions of prostate cancer. Point mutations were retrieved from the COSMIC database and then subjected to computational analysis. Recurrent non-coding mutations were identified in MIR100HG, MIR4432, and MIR2054. These statistically significant mutations were mapped to prostate cancer cell-line specific TFBS, revealing their presence within recurrent mutational hotspots. Importantly, a 10,000-bootstrap randomization model ($q < 0.05$) confirmed that these hotspots are unlikely to be random. The loci identified here represent strong candidate regulatory regions that warrant experimental validation through approaches like luciferase reporter assays, and CRISPR-based editing, as well as clinical cohort studies to establish their functional and causal relevance. Overall, this framework highlights coordinated disruptions across transcriptional and post-transcriptional networks, offering a novel computational strategy to distinguish biologically meaningful regulatory loci from background genetic noise. By doing so, it provides a foundation for future biomarker discovery and the development of targeted interventions in prostate cancer.

Keywords: Cell lines, microRNAs, prostate cancer, recurrent non-coding mutations, transcription factor binding sites

INTRODUCTION

Prostate cancer is the 2nd most prevalent cancer among men worldwide and 5th leading cause of cancer related to mortality, with approximately 1.4 million new cases reported annually (Akhtar *et al.*, 2023). It is a complex and multifactorial disease influenced by genetic, environmental, and lifestyle factors, with genetic mutations serving as key drivers of its initiation and progression (Ren *et al.*, 2023; Gammall and Lai, 2024). While much of the research has traditionally focused on coding mutations within protein-coding genes, the non-coding genome is now recognized as a critical contributor to cancer biology (Woo *et al.*, 2024). Mutations in non-coding regions—particularly in enhancers and transcription factor binding sites (TFBS)—can profoundly alter gene expression and cellular function (Handler *et al.*, 2024). A comprehensive analysis of genetic landscape of prostate cancer must therefore integrate both coding and non-coding mutations. TFBS are especially important, as they represent genomic regions where transcription factors—the central regulators of gene expression—bind to orchestrate complex

regulatory programs. Mutations within these sites disrupt transcriptional control, leading to dysregulated gene expression networks that contribute to prostate cancer progression (Peng *et al.,* 2024).

MicroRNAs (miRNAs) are small non-coding RNA molecules that act as essential regulators of gene expression, participating in processes like apoptosis, differentiation, and cell division. Dysregulation of miRNA expression has consistently been linked to the growth and progression prostate cancer (Gujrati *et al.*, 2023). Emerging evidences highlight the broader role of non-coding RNAs (ncRNAs)—including both miRNAs and long non-coding RNAs (lncRNAs)—in prostate cancer pathogenesis, diagnosis, and therapy. these ncRNAs influence tumor progression through signalling pathways such as Wnt, underscoring the biological relevance of miRNA dysregulation in this malignancy (Bu *et al.*, 2023; Taheri *et al.*, 2023; Zhou *et al.*, 2023). Several miRNAs have been identified as potential diagnostic and prognostic biomarkers. For instance, miR-21, miR-141, and miR-375, are strongly associated with prostate cancer progression (Goztepe and Eroglu, 2024). Conversely, miR-34a functions as tumor suppressor by inhibiting cell growth and regulating the p^{53} pathway (Abubakar and Rehman, 2024). MiRNAs also regulate epithelial-mesenchymal transition (EMT), a key process in cancer metastasis. Members of miR-200 family, particularly miR-200c, reportedly inhibit EMT in prostate cancer (Xie *et al.*, 2024). Furthermore, miRNAs such as miR-21 and miR-125b modulate androgen receptor (AR) expression and activity, a central axis in prostate cancer biology (Singh *et al.*, 2024).

Genetic alterations in prostate cancer, involve both coding mutations such as TP53, PTEN, and AR, and non-coding regulatory elements such as enhancers and TFBS which drive oncogenic transcriptional programs (Cussenot *et al.*, 2023; Singh *et al.*, 2024; Kulac *et al.*, 2024; Iñiguez-Muñoz *et al.*, 2024). While genomic profiling has traditionally focused on protein-coding drivers, the non-coding genome contains structural variations that profoundly influence gene expression and cellular functions relevant to the cancer. These non-coding elements include transcription factors and miRNAs that regulate each other in a tightly coupled manner i.e. transcription factors regulate miRNA transcription, while miRNAs post-transcriptionally modulate transcription factors. Transcription factors and miRNAs thus form tightly interconnected regulatory networks in prostate cancer; and mutations occurring in either TFBS or miRNA regulatory regions collectively disrupt these feedback loops, leading to abnormal gene expression. The somatic mutations in TFBS affect transcription of protein-coding genes as well as miRNA expression, providing evidence that mutations in regulatory DNA cascade trans-effects on gene regulatory networks, including miRNAs (Castro-Mondragon *et al.*, 2022). The integrative approaches combining whole-genome mutation data, regulatory annotations, and transcriptomics improve detection of candidate-driver mutations (Nourbakhsh *et al*., 2024), supporting the rationale for examining non-coding mutations within both TFBS and miRNA-associated regions to identify potential regulatory hotspots.

Most computational models have focused on the changes in either TFBS or miRNA, but none on the changes in both. This study proposed an integrative framework analysing the recurrent mutations in both TFBS or miRNA regions. Unlike prior models restricted to one category, it identifies recurrent non-coding mutations as candidate regulatory hotspots. Recurrence is treated as statistical evidence of candidate regulatory relevance; prioritizing loci for future experimental and clinical validation, without attributing direct oncogenic causality.

MATERIALS AND METHODS

A four-part investigative strategy employed in this study included: 1) stringent filtering of localized non-coding somatic variants from multiple sources; 2) statistical aggregation to identify the high-frequency recurrent genomic coordinates; 3) spatial mapping across the extended miRNA genomic regions; and 4) bootstrap randomization to separate true non-coding hotspots from background mutational noise. To preserve the biological distinction, a consistent (recurrent) mutation was

considered to be a somatic point mutation arising in the same genomic coordinate in five or more different tumor samples or cell lines within the cohort. A driver mutation, on the other hand, suggests that the mutation is definitely known to be functional in driving oncogenesis. The loci identified in this study are not necessarily causal; given the computational and statistical scale of the study, they are presented as candidate regulatory hotspots that should serve as a roadmap for further laboratory-based validation.

Data collection

The mutation data in non-coding variants were retrieved from the COSMIC database (dataset version v98, released 2024) as it offers extensive data on somatic mutations across a variety of cancer types (Tang *et al.,* 2023). To obtain mutations, the dataset was filtered for the following criteria: cancer type, 'prostate'; sample types, tumor tissue and prostate cancer cell lines only; and mutation category, point mutations within non-coding regions. The resulting file contained 2,49,743 prostate cancer samples, which were further processed using Linux and Bedtools utilities (v2.30.0) for sorting and merging. Bedtools encompasses a range of commands that are well-suited to manage vast genomic datasets. Fig. 1 illustrates the workflow for mutational analysis of non-coding variants in prostate cancer.

Identification of consistent mutations in non-coding variants

Consistent mutations are somatic point mutations occurring at the same genomic location in two or more samples from the same cluster (Wakeley *et al.,* 2023). Identifying consistent mutations in non-

consistent mutations are somatic point mutations occurring at the same genomic location in two or more samples from the same cluster (Wakeley *et al.,* 2023). The identification of consistent mutations in non-coding regions in cancers is a critical step in uncovering the molecular mechanisms underlying the disease (Sethi and Shar, 2023). These consistent mutations recur at specific genomic locations within a given cancer type. The next step, therefore, was to identify mutations consistently observed across prostate cancer samples using the Bedtools intersect and overlap commands. In this study, a mutation was considered "consistent" only if its recurrence was ≥5. This threshold reflects these mutations that can be identified statistically but without being functionally proven to be the drivers of onco-genesis. The terms 'consistent' and 'recurrent' are interchangeable throughout this manuscript and do not imply "driver" status.

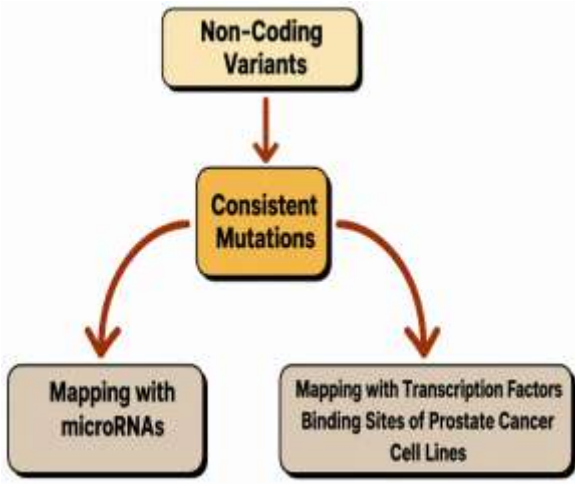

Fig. 1: Schematic illustration of workflow showing the identification of consistent non-coding mutations and their mapping with miRNAs and transcription factor binding sites

Mapping consistent non-coding mutations with miRNAs and their functional annotation

The genomic coordinates (chromosome, start position, end position) of all miRNAs were downloaded from the Ensembl Genome Browser (release 109, GRCh38) (*https://www.ensembl.org/*). The miRNA genomic locations were extended by up to 100 kb to assess potential cis-regulatory interactions, since approximately 60-70% of enhancer–gene interactions reportedly occur within this distance (Gschwind *et al.,* 2023; Vattathil *et al.,* 2025). To enhance biological specificity and reduce the inclusion of non-functional intergenic background, all variants identified within the extended ±100 kb regions were filtered using regulatory annotations. Publicly available miRNA target prediction tools, such as KEGG pathways and Gene Ontology (GO) categories, were used for functional enrichment analysis of the most recurrently mutated miRNAs.

Identification of consistent non-coding mutations in regulatory regions

Regulatory regions play a crucial role in controlling gene expression and includes promoters, enhancers, silencers, transcription factor binding sites, insulators, and cis-regulatory modules (CRMs) (Yuan *et al.*, 2025). Mutations within these regions can disrupt transcription binding process, resulting in abnormal—either excessive or suppressed— gene expression. This study, therefore, identified mutations within TFBS of prostate cancer cell lines. Prostate cancer cell lines were first identified, followed by their corresponding transcription factors, as shown in Fig. 2. Binding-site data for these transcription factors were retrieved from UCSC ENCODE (2023 ENCODE cCRE and TF ChIP-seq tracks). Consistent non-coding mutations were then mapped to these TFBS to determine whether the point mutations occurred within regulatory regions.

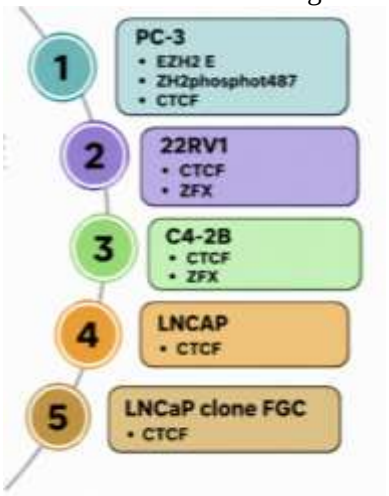

Fig. 2: Prostate cancer cell lines and their associated transcription factors used for TFBS mapping, including PC-3, 22RV1, C4-2B, LNCaP, and LNCaP clone FGC, identified from the CHIP-seq experiment

Overlap analysis between TFBS regions and mutated miRNAs

To investigate potential combinatorial effects, an overlap analysis was performed using the Bedtools intersect and overlap commands on the genomic coordinates of TFBS and consistently mutated miRNA regions.

Statistical significance of acquired results

Randomization, a powerful technique, was used for various statistical analysis. In this study, a bootstrap randomization test (10,000 iterations) was performed to determine the significance of identified mutations. Mutation coordinates were randomly shuffled across the non-coding genome while preserving the original mutation count per iteration. The expected distribution of recurrence was computed and compared to the observed recurrence, and the *p*-value calculated as the proportion of randomized iterations where recurrence ≥ was observed. Multiple-testing correction was employed using Benjamini–Hochberg False Discovery Rate (FDR) method. Mutations were considered statistically significant only when the FDR-adjusted *p*-value (*q*-value) was <0.05.

RESULTS AND DISCUSSION

Identification of consistent mutations in non-coding variants

The information on genomic locations of consistent mutations in non-coding regions is shown in Fig. 3. Each row represents the number of consistent mutations observed at a specific genomic location. The highest frequency of consistent mutations was observed on chromosome 3 (Chr3), with 189 occurrences, followed by Chr9, Chr8, Chr19, and Chr12. The strong enrichment of mutational events on Chr3 suggested that this region is highly unstable in prostate somatic tissue, possibly reflecting region-specific selective pressures on the non-coding genome; however, this observation alone does not establish the underlying mechanism.

Mapping consistent non-coding mutations with miRNAs

The consistent non-coding mutations identified within miRNA genomic locations, extended by 100 kb, in prostate cancer are depicted in Table 1. Each row reports the frequency of consistent mutations

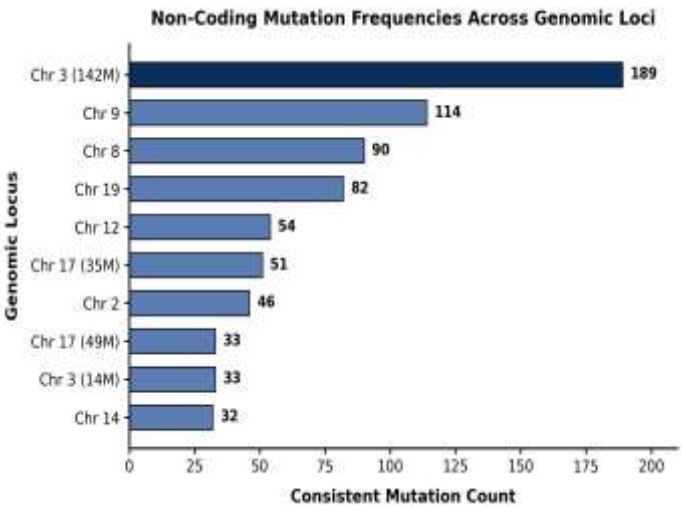

Fig. 3: The consistent non-coding point mutations, reported at the ten highest-recurrence sites (Table 1), are plotted on the chromosomes, with dominant hotspot being chromosome 3 (n = 189)

Table 1: Consistent non-coding mutations found in miRNA 100 kb extended genomic regions

Genomic coordinates	MiRNA	Consistency
11: 121829063-122030463	MIR100HG	91
2: 60114360-60315380	MIR4432	90
4:125418677-125620859	MIR2054	88
4:11041255-11244027	MIR6886	85
15:37767306-37968806	MIR8063	84
3:102567318-102769538	MIR548A3	84
14:23337378-23539005	MIR2088	79
14:29044097-29245522	MIR548AI	79
18:37498286-37699167	MIR4318	79
14:23342404-23543714	MIR208B	79

(MIR4432/MIR2054, 90 occurrences), both of which are proven regulators of proliferation and resistance to treatment in chronic prostate cancer. The pathway associations in this study were not generated with the matched expression data, so they should be interpreted as literature-based hypotheses and not as a result of this study.

The high recurrence observed in MIR100HG, MIR4432, and MIR2054 suggests potential disruption of regulatory elements involved in prostate cancer-specific pathways such as PI3K-Akt signalling and androgen receptor regulation. Dai *et al.* (2025) reported that MIR100HG promotes cell proliferation in (consistency) of a specific chromosomal region associated with a given miRNA (MIR). MIR100HG exhibited the highest consistency, with 91 occurrences, suggesting its potential key role in prostate cancer. High consistency was also observed for MIR4432, MIR2054, and MIR6886, with 90, 88, and 85 occurrences, respectively. The functional annotation of these highly mutated miRNAs revealed that MIR100HG targets were enriched in the PI3K-Akt signalling pathway, a pathway well known for its role in prostate cancer progression. Similarly, MIR4432 targets were found associated with androgen receptor signalling and apoptosis regulation. These findings indicate a possible association of the identified recurrent miRNA mutations with the key oncogenic pathways involved in prostate cancer, but such an association at an expression level and mechanistic levels need confirmation. The pathway links associated with the three most frequently mutated miRNA loci are summarized in Fig. 4. These sites were consistent across the genome and are known to be associated with PI3K-Akt signalling cascade (MIR100HG, 91 occurrences) and androgen receptor (AR) signalling, and apoptosis regulation

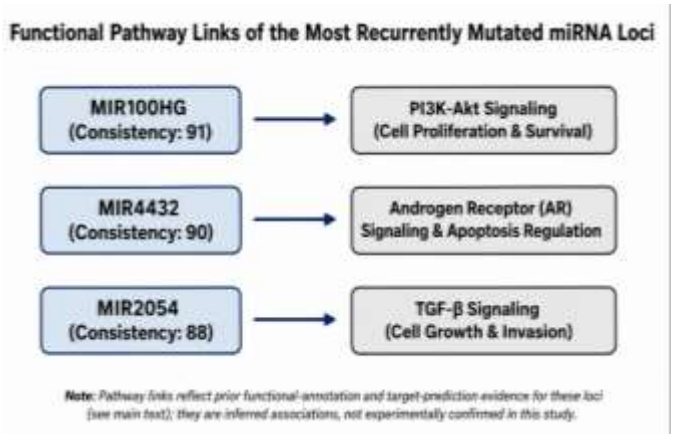

Fig. 4: Schematic mapping of the most frequently mutated miRNA loci (MIR100HG, MIR4432, and MIR2054) onto previously reported downstream oncogenic pathways including PI3K-Akt signalling and regulation of androgen receptor/apoptosis

breast cancer. The high consistency of this miRNA in this study suggests its potential role in prostate cancer as well. The recurrence of mutations across other miRNAs emphasized the importance of these regulatory elements in the pathogenesis of prostate cancer. This concept is widely supported by the studies demonstrating that non-coding mutations in the regulatory plexus of the key oncogenes, such as FOXA1, can affect transcription factor binding and enhancer activity, thereby driving prostate cancer progression (Morova *et al.*, 2020). A notable example of this mechanism is the functional validation of non-coding enhancer GH221030351 as a driver of metastatic prostate cancer growth (Wang *et al.*, 2023). The functional annotation of recurrently mutated miRNAs revealed enrichment in keyprostate cancer-related pathways, such as androgen receptor regulation and PI3K-Akt signalling pathway (He *et al.*, 2022). These associations highlight the biological relevance of the acquired results.

Identification of consistent non-coding mutations in regulatory regions

Consistent non-coding mutations identified within TFBS of prostate cancer cell lines is given in Table 2. Each row reports the frequency of consistent mutations (consistency) identified at specific genomic locations and the number of transcription factors binding to those locations. Although mutation consistency was high on chromosomes 3 and 9 (189 and 114, respectively), these regions showed no transcription factor binding. However, the regions with lower consistency displayed a higher number of transcription factor binding, suggesting a potential regulatory role. This identification of highly consistent mutations, despite limited transcription factor binding, raises concerns about these regulatory regions. A recent mechanistic study offers a possible explanation, which showed the binding of androgen receptor (AR), as a crucial transcription factor in prostate cancer, physically inhibited the DNA repair machinery at its binding sites, resulting in remarkably elevated local mutation rate specifically at its binding sites (Saha *et al.*, 2024). Therefore, a more thorough investigation of these transcription factors and their interactions with mutant non-coding areas is required.

Table 2: Consistent non-coding mutations identified within transcription factor binding sites of prostate cancer cell lines

Genomic coordinates	Consistency	No. of transcription factors binding
3:142449486-142449486	189	0
9:132897139-132897139	114	0
19: 11060163-11060163	82	0
2:29193928:29193928	46	0
14:100285792-100285792	12	2
19:58519652-58519652	8	3
22:21629335-21629335	6	2
16:30902217-30902217	5	3
19:17282904-17282904	5	2
2:70994919-70994919	5	1

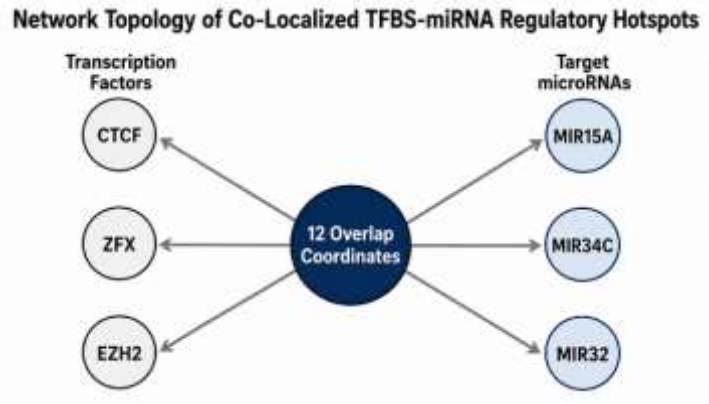

Fig. 5: Candidate combinatorial regulatory hotspots: Network topology of 12 co-localized genomic coordinates that are mutated in both transcription factor binding sites (CTCF, ZFX, EZH2) and target miRNA sites (MIR15A, MIR34C, MIR32)

Overlap analysis between TFBS and mutated miRNA regions

The overlap analysis revealed co-localization of consistent mutations in a subset of regions (n = 12) in both TFBS and miRNA regulatory regions. MIR15A showed highest overlap among all the regions examined. The expression of MIR15A was found influenced by transcription factors in prostate cancer, placing it among the deregulated miRNA clusters implicated in this disease (Xiang *et al.*, 2025). MIR34C was down-regulated in prostate cancer via USP2a-mediated Myc activation,

implying that the transcription factor 'Myc' regulates the miRNA locus (Ahmadi *et al.*, 2021). MIR32 is likewise regulated in prostate cancer, often under androgen/AR control (Fatemeh *et al.*, 2024). These overlapping regions may represent regulatory hotspots that disrupt both transcription factor binding and miRNA transcription simultaneously, producing a combined effect on gene dysregulation (Fig. 5). TFBS-associated recurrent mutations, particularly within CTCF-, ZFX-, and EZH2-binding loci, highlight possible interference with the transcriptional control of prostate-specific genes. These overlapping patterns imply coordinated disruption of both transcriptional and post-transcriptional regulations in prostate cancer. Many lesser-known miRNAs with high consistencies were found among those co-regulated by different transcription factors, such as MIR1302 clusters and the MIR548 family. The observed overlap between TFBS and mutated miRNA regions suggests the existence of combinatorial regulatory disruption, highlighting these regions to be potential molecular 'hotspots'. This is consistent with the findings at established driver loci, such as FOXA1 regulatory plexus, where groups of non-coding mutations act together to disrupt gene regulation (Zhou *et al.*, 2020).

Statistical significance of identified mutations

The statistically significant consistent non-coding mutations within TFBS in prostate cancer cell lines (Table 3) highlight their potential role in the regulatory landscape of prostate cancer. The regions with high mutation consistency showed binding of only one transcription factor. The

Table 3: Statistically significant mutations identified in non-coding variants in prostate cancer

Genomic coordinates	Consistency	Names of TFBS	*p*-value*
17:49619064-49619064	33	CTCF	0.014499
3:14159852-14159852	33	ZFX	0.006556
2:23360891-23360891	30	ZFX	0.033324
18:35241018-35241018	30	CTCF	0.030773
17:62423853-62423853	30	ZFX	0.043906
19:1592756-1592756	30	EZH2	0.029486
2:85455834-85455834	12	ZFX, CTCF	0.006834
17:81982516-81982516	10	ZFX, CTCF	0.006726
14:100076870-100076870	5	ZFX, CTCF, EZH2	0.049799
16:30902217-30902217	5	CTCF	0.041953

TFBS: Transcription factor binding sites; **p*-values were computed by a 10000-iteration bootstrap randomization test as the proportion of randomized iterations with recurrence ≥ the observed value; mutations were considered statistically significant at $p < 0.05$, with a Benjamini–Hochberg FDR-adjusted $q < 0.05$ applied for multiple-testing correction.

associated p-values (all < 0.05; Table 3) indicated that these mutations are statistically significant. These results suggest that the observed consistency and transcription of binding factors are unlikely to occur by chance alone (Fig. 6).

Summary of key findings, limitations, and future directions

Prostate cancer is one of the most prevalent malignancies in males worldwide. The biology of prostate cancer is complex and needs thorough insight. This study presented a novel integrative computational framework to identify recurrent non-coding mutations in prostate cancer, with focus on miRNAs and TFBS. This methodology aligns with the growing recognition of importance of non-coding regulatory elements in cancer and the development of sophisticated frameworks that integrate computational approaches with functional validation like that of Woo *et al*. (2024), who combined computational prediction with luciferase and CRISPR-based functional assays to validate non-coding regulatory regions driving metastatic prostate cancer. Unlike previous studies that have predominantly investigated either mutations in the coding region or the role of miRNAs in the development and progression of prostate cancer (Rana *et al.*, 2022; Schitcu *et al.*, 2022), this study focuses on recurrent and statistically validated point mutations within non-coding regulatory elements. The analysis of miRNA mutations revealed a consistent pattern across various samples, suggesting a significant impact

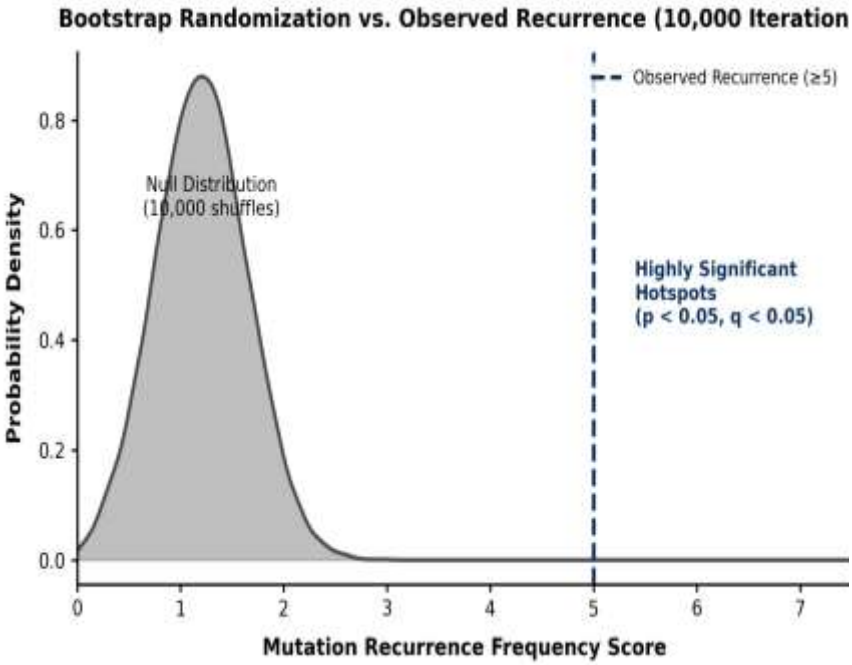

Fig. 6: Mutational significance evaluated by bootstrap randomization analysis (10,000 iterations). The shaded curve represents the null distribution of the expected number of non-coding mutations recurring as a result of shuffling mutation coordinates throughout the genome, and the dashed vertical line indicates the observed recurrence threshold (≥ 5), indicating that the identified hotspots are unlikely to occur by chance ($p < 0.05$, $q < 0.05$, FDR-corrected).

on miRNA functionality and target interactions, potentially leading to alterations in the regulation of gene expression. Further investigations into the specific mutations and their effects on target miRNA binding sites may be crucial to understand the underlying mechanisms and implications of this disease. The single most recurrent locus identified in this study reached a count of 189 occurrences, which describes its recurrence rather than its casual status. The significant levels of mutations provide crucial insights into the robustness and validity of the obtained results. There are significant clinical implications of consistent non-coding mutations identified within miRNA and TFBS regions. The convergence of miRNA and TFBS mutation hotspots provides insight into the existence of regulatory hubs that may be critical for prostate cancer progression. Since early detection of prostate cancer is a challenging task, so these mutations may serve as potential biomarkers for early prostate diagnosis and treatment. The integration of this genomic information with clinical data might facilitate the development of more personalized therapies and targeted therapeutic approaches.

This study is primarily computational and statistical in nature, which limits the conclusions that can be drawn in several specific ways. Most importantly, there is no functional validation: no luciferase reporter assays, CRISPR-based editing, or other wet laboratory experiments are presented to demonstrate that the identified mutations affect the transcription factor binding affinity, miRNAs processing, or both. Consequently, the biological relevance of the reported loci remains speculative until tested experimentally. Second, this study relies exclusively on public repositories (COSMIC, Ensembl, UCSC ENCODE) and prostate cancer cell lines; no patient-level clinical data or diverse patient cohorts were analysed. It is therefore not possible to relate these mutations to disease stage, prognosis or response to therapy, which limits translational potential. Third, the focus was restricted to TFBS and miRNA regions, without examining other non-coding regulatory elements such as enhancers, silencers, or long non-coding RNAs (lncRNAs), which may also be important in prostate cancer biology. Furthermore, associations between miRNAs and pathway-level events (e.g., recurrently mutated miRNAs and PI3K-Akt or androgen receptor signalling) were inferred from existing literature rather than expression data generated in this study. Finally, although the identified loci are proposed as candidate biomarkers, they have not been validated in clinically relevant systems (e.g., tissue or liquid biopsy), leaving the translation pathway from the computational hotspot to validated biomarker incomplete. Taken together, these limitations suggest that the finding should be

considered as a prioritized, statistically supported list of potential regulatory and/or clinical biomarkers, rather than definitive functional or clinical markers. To advance this framework, two-phase validation pathway is recommended. First, *in vitro* functional assays such as dual-luciferase reporter experiments using synthetic vectors with wild-type and mutated promoter/enhancer regions, and targeted CRISPR-Cas9 base editing in prostate cancer cell lines (PC-3, LNCaP), could clarify phenotypic effects on cell motility and proliferation. Second, clinical cohort screening using publicly available patient-level datasets (TCGA, ICGC prostate adenocarcinoma) could correlate these non-coding mutations with matched RNA sequencing and clinical outcome data, including disease stage, biochemical recurrence-free survival, and therapeutic response. Together, these steps would transform the current hypothesis-generating computational pipelines into a biomarker discovery platform for prostate cancer with functional and clinical validation.

Conclusion: The integrative approach employed in this study identified statistically significant, highly frequent non-coding mutational hotspots in prostate cancer genome. Notably, three miRNA's—MIR100HG, MIR4432 and MIR2054—together with recurrent mutations in the binding domains of CTCF and ZFX, emerged as the potential candidate regulatory sites. These hotspots warrant further investigations to understand their role in prostate cancer progression. As this study relied on a computational and/or predictive approach, therefore identifying the functionally validated, and causally involved driver mutations within these loci is the next step. This integrative framework provides a new computational starting point for future biomarker discovery pipelines, non-invasive liquid biopsy development, and personalized diagnostic and therapeutic strategies in prostate cancer.

Acknowledgments: The authors acknowledge support received from the Department of Biomedical Engineering, NED University of Engineering & Technology, Pakistan.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

Abubakar, M. and Rehman, B. 2024. Roles of mutant TP53 gene in cancer development and progression. *Proceedings of Anticancer Research*, **8**(5): 165-181.

Ahmadi, S.E., Rahimi, S., Zarandi, B., Chegeni, R. and Safa, M. 2021. MYC: A multipurpose oncogene with prognostic and therapeutic implications in blood malignancies. *Journal of Hematology & Oncology*, **14**(1): 121. [*https://doi.org/10.1186/s13045-021-01111-4*].

Akhtar, S., Hassan, F., Ahmad, S., El-Affendi, M.A. and Khan, M.I. 2023. The prevalence of prostate cancer in Pakistan: A systematic review and meta-analysis. *Heliyon*, **9**(10): e20350. [*https://doi.org/10.1016/j.heliyon.2023.e20350*].

Bu, T., Li, L. and Tian, J. 2023. Unlocking the role of non-coding RNAs in prostate cancer progression: Exploring the interplay with the Wnt signaling pathway. *Frontiers in Pharmacology*, **14**: 1269233. [*https://doi.org/10.3389/fphar.2023.1269233*].

Castro-Mondragon, J.A., Aure, M.R., Lingjærde, O.C., Langerød, A., Martens, J.W.M., Børresen-Dale, A.L., *et al.*, 2022. Cis-regulatory mutations associate with transcriptional and post-transcriptional deregulation of gene regulatory programs in cancers. *Nucleic Acids Research*, **50**(21): 12131-12148.

Cussenot, O., Cancel-Tassin, G., Rao, S.R., Woodcock, D.J., Lamb, A.D., Mills, I.G., *et al.*, 2023. Aligning germline and somatic mutations in prostate cancer. Are genetics changing practice? *BJU International*, **132**(5): 472-484.

Dai, M., Shi, Y., Zhao, H., Hu, Y., Cong, X., Yu, B., *et al.*, 2025. A novel MIR100HG transcript enhances tumorigenesis by inducing BCLAF1-mediated alternative splicing in colorectal cancer. *Cell Communication and Signalling*, **23**(1): 328. [*https://doi.org/10.1186/s12964-025-02280-2*].

Fatemeh, S., Mahboobeh, Z., Khadijeh, A., Amirhossein, M.K. and Pegah, M. 2024. An *in silico* study to determine susceptibility to cancer by evaluating the coding and non-coding non-synonymous single nucleotide variants in the SOCS3 gene. *Journal of Biomolecular Structure and Dynamics*, **42**(16): 8281-8292.

Gammall, J. and Lai, A.G. 2024. Prognostic determinants in cancer survival: A multidimensional evaluation of clinical and genetic factors across 10 cancer types in the participants of Genomics England's 100,000 Genomes Project. *Discover Oncology*, **15**(1): 448. [*https://doi.org/10.1007/s12672-024-01310-8*].

Goztepe, M. and Eroglu, O. 2024. Research of the unrecognised functions of miR-375 in prostate cancer cells. *Cellular and Molecular Biology*, **70**(3): 212-218.

Gschwind, A.R., Mualim, K.S., Karbalayghareh, A., Sheth, M.U., Dey, K.K., Jagoda, E., *et al.*, 2023. An encyclopedia of enhancer-gene regulatory interactions in the human genome. *bioRxiv*, 2023.11.09.563812 (Preprint). [*https://doi.org/10.1101/2023.11.09.563812*].

Gujrati, H., Ha, S. and Wang, B.D. 2023. Deregulated microRNAs involved in prostate cancer aggressiveness and treatment resistance mechanisms. *Cancers (Basel)*, **15**(12): 3140. [*https://doi.org/10.3390/cancers15123140*].

Handler, J.S., Li, Z., Dveirin, R.K., Fang, W., Goodarzi, H., Fertig, E.J., *et al.*, 2024. Identifying a gene signature of metastatic potential by linking the pre-metastatic state to ultimate metastatic fate. *bioRxiv*, 2024.08.14.607813 (Preprint) [*https://doi.org/10.1101/2024.08.14.607813*].

He, Y., Xu, W., Xiao, Y.T., Huang, H., Gu, D. and Ren, S. 2022. Targeting signaling pathways in prostate cancer: Mechanisms and clinical trials. *Signal Transduction and Targeted Therapy*, **7**(1): 198. [*https://doi.org/10.1038/s41392-022-01042-7*].

Iñiguez-Muñoz, S., Llinàs-Arias, P., Ensenyat-Mendez, M., Bedoya-López, A.F., Orozco, J.I.J., Cortés, J., *et al.*, 2024. Hidden secrets of the cancer genome: Unlocking the impact of non-coding mutations in gene regulatory elements. *Cellular and Molecular Life Sciences*, **81**(1): 274. [*https://doi.org/10.1007/s00018-024-05314-z*].

Kulac, I., Roudier, M.P. and Haffner, M.C. 2024. Molecular pathology of prostate cancer. *Clinics in Laboratory Medicine*, **44**(2): 161-180.

Morova, T., McNeill, D.R., Lallous, N., Gönen, M., Dalal, K., Wilson, D.M., *et al.*, 2020. Androgen receptor-binding sites are highly mutated in prostate cancer. *Nature Communications*, **11**(1): 832. [*https://doi.org/10.1038/s41467-020-14644-y*].

Nourbakhsh, M., Degn, K., Saksager, A., Tiberti, M. and Papaleo, E. 2024. Prediction of cancer driver genes and mutations: The potential of integrative computational frameworks. *Briefings in Bioinformatics*, **25**(2): [*bbad519. https://doi.org/10.1093/bib/bbad519*].

Peng, Y., Song, W., Teif, V.B., Ovcharenko, I., Landsman, D. and Panchenko, A.R. 2024. Detection of new pioneer transcription factors as cell-type-specific nucleosome binders. *eLife*, **12**: RP88936. [*https://doi.org/10.7554/eLife.88936*].

Rana, S., Valbuena, G.N., Curry, E., Bevan, C.L. and Keun, H.C. 2022. MicroRNAs as biomarkers for prostate cancer prognosis: A systematic review and a systematic reanalysis of public data. *British Journal of Cancer*, **126**(3): 502-513.

Ren, S., Li, J., Dorado, J., Sierra, A., González-Díaz, H., Duardo, A., *et al.*, 2023. From molecular mechanisms of prostate cancer to translational applications: Based on multi-omics fusion analysis and intelligent medicine. *Health Information Science and Systems*, **12**(1): 6. [*https://doi.org/10.1007/s13755-023-00264-5*].

Saha, D., Dang, H.X., Zhang, M., Quigley, D.A., Feng, F.Y. and Maher, C.A. 2024. Single cell-transcriptomic analysis informs the lncRNA landscape in metastatic castration-resistant prostate cancer. *NPJ Genomic Medicine*, **9**(1): 14. [*https://doi.org/10.1038/s41525-024-00401-3*].

Schitcu, V.H., Raduly, L., Nutu, A., Zanoaga, O., Ciocan, C., Munteanu, V.C., *et al.*, 2022. MicroRNA dysregulation in prostate cancer. *Pharmacogenomics and Personalized Medicine*, **15**: 177-193.

Sethi, A.A. and Shar, N.A. 2023. Identification of liver cancer driver mutations from COSMIC data. *International Journal of Cancer Management*, **16**(1): e131281. [*https://doi.org/10.5812/ijcm-131281*].

Singh, V.K., Rajak, N., Singh, Y., Singh, A.K., Giri, R. and Garg, N. 2024. Role of microRNA-21 in prostate cancer progression and metastasis: Molecular mechanisms to therapeutic targets. *Annals of Surgical Oncology*, **31**(7): 4795-4808.

Taheri, M., Badrlou, E., Hussen, B.M., Kashi, A.H., Ghafouri-Fard, S. and Baniahmad, A. 2023. Importance of long non-coding RNAs in the pathogenesis, diagnosis, and treatment of prostate cancer. *Frontiers in Oncology*, **13**: 1123101. [*https://doi.org/10.3389/fonc.2023.1123101*].

Tang, Q., Zuo, W., Wan, C., Xiong, S., Xu, C., Yuan, C., *et al.*, 2023. Comprehensive genomic profiling of upper tract urothelial carcinoma and urothelial carcinoma of the bladder identifies distinct molecular characterizations with potential implications for targeted therapy and immunotherapy. *Frontiers in Immunology*, **13**: 1097730. [*https://doi.org/10.3389/fimmu.2022.1097730*].

Vattathil, S.M., Gerasimov, E.S., Canon, S.M., Lori, A., Tan, S.S.M., Kim, P.J., *et al.*, 2025. Mapping the microRNA landscape in the older adult brain and its genetic contribution to neuropsychiatric conditions. *Nature Aging*, **5**(2): 306-319.

Wakeley, J., Fan, W.L., Koch, E. and Sunyaev, S. 2023. Recurrent mutation in the ancestry of a rare variant. *Genetics*, **224**(3): iyad049. [*https://doi.org/10.1093/genetics/iyad049*].

Wang, Z., Luo, M., Liang, Q., Zhao, K., Hu, Y., Wang, W., *et al.*, 2023. Landscape of enhancer disruption and functional screen in melanoma cells. *Genome Biology*, **24**(1): 248. [*https://doi.org/10.1186/s13059-023-03087-5*].

Woo, B.J., Moussavi-Baygi, R., Karner, H., Karimzadeh, M., Yousefi, H., Lee, S., *et al.*, 2024. Integrative identification of non-coding regulatory regions driving metastatic prostate cancer. *Cell Reports*, **43**(9): 114764. [*https://doi.org/10.1016/j.celrep.2024.114764*].

Xiang, Z., Lin, T., Ling, J., Xu, Z., Huang, R., and Hu, H. 2025. MiRNA expression profiling and clinical implications in prostate cancer across various stages. *Scientific Reports*, **15**(1): 7771. [*https://doi.org/10.1038/s41598-025-92091-9*].

Xie, H., Chen, J., Ma, Z., Gao, Y., Zeng, J., Chen, Y., *et al.*, 2024. PrLZ regulates EMT and invasion in prostate cancer via the TGF-β1/p-Smad2/miR-200 family/ZEB1 axis. *Prostate*, **84**(4): 317-328.

Yuan, S., Ni, P. and Su, Z. 2025. Prediction of target genes and functional types of cis-regulatory modules in the human genome reveals their distinct properties. *BMC Biology*, **23**(1): 211. [*https://doi.org/10.1186/s12915-025-02313-9*].

Zhou, H., Hao, X., Zhang, P. and He, S. 2023. Noncoding RNA mutations in cancer. *Wiley Interdisciplinary Reviews: RNA*, **14**(6): e1812. [*https://doi.org/10.1002/wrna.1812*].

Zhou, S., Hawley, J.R., Soares, F., Grillo, G., Teng, M., Madani Tonekaboni, S.A., *et al.*, 2020. Noncoding mutations target cis-regulatory elements of the FOXA1 plexus in prostate cancer. *Nature Communications*, **11**(1): 441. [*https://doi.org/10.1038/s41467-020-14318-9/*].