

Comprehensive Analysis of *HOXB13* Expression, Prognosis, and Pathogenic Variants in Prostate Adenocarcinoma Across Ethnically Diverse Populations

Lalu Muhammad Irham^{1,2,*}, Wirawan Adikusuma³, Syarifatul Mufidah¹, Rahmat Dani Satria^{4,5}, Eko Mugiyo⁶, Indra Jaya^{7,8,9}, Darmawi Darmawi¹⁰, Petrina Theda Philothra¹¹, Desti Rahmawati¹, Siska Widiastuti¹, Rista Yulianti¹, Putri Permata Suka Admanegara¹, Brilliant Citra Wirashada¹², Baik Heni Risprawati¹³, Sabiah Khairi¹⁴, Rockie Chong¹⁵, Ilker Ates¹⁶

¹Faculty of Pharmacy, Universitas Ahmad Dahlan, Yogyakarta, Indonesia; ² Faculty of Pharmacy, Silpakorn University, Nakhon Pathom, 73000, Thailand; ³Research Centre for Pharmaceutical Ingredients and Traditional Medicine, National Research and Innovation Agency (BRIN), South Tangerang, Indonesia; ⁴Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia; ⁵Integrated Laboratory Installation, Dr. Sardjito General Hospital, Yogyakarta, 55281, Indonesia; ⁶Department of Pharmacy, Faculty of Health Science, University of Muhammadiyah Pekalongan, Pekalongan, Indonesia; ⁷Urology Division, Surgery Department, Arifin Achmad General Regional Hospital; ⁸Faculty of Medicine Universitas Riau, Pekanbaru, Riau, Indonesia; ⁹ Faculty of Medicine Universitas Padjadjaran, Bandung, West Java, Indonesia; ¹⁰Department of Histology, Faculty of Medicine, Universitas Riau, Diponegoro St No.1, Pekanbaru, Indonesia; ¹¹ Departement of Rehabilitation Medicine, General Hospital Yogyakarta City, Yogyakarta, Indonesia; ¹²Department of Surgery, Faculty of Medicine, Universitas Muhammadiyah Surabaya, Surabaya, Indonesia; ¹³ Institut Kesehatan YARSI Mataram, Mataram, Indonesia; ¹⁴Faculty of medicine and health science, Universitas Islam Negeri Mataram, Mataram, 83116, Indonesia ¹⁵Department of Chemistry and Biochemistry, University of California, Los Angeles, USA; ¹⁶Faculty of Pharmacy, Department of Toxicology, Ankara University, Yenimahalle, Ankara 06560, Turkey

Received: October 24, 2025; Revised: February 1, 2026; Accepted: February 23, 2026

Abstract

Background: *HOXB13* is a transcription factor known to play a critical role in prostate development and Prostate Adenocarcinoma (PRAD). However, a systematic evaluation of its expression patterns, prognostic relevance, and pathogenic variants across tissues, cancer types, and populations remains limited. This study hypothesized that *HOXB13* exhibits prostate-specific overexpression in PRAD with prognostic significance across ethnic groups and aimed to comprehensively characterize its expression profiles, prognostic value, and functionally pathogenic variants. **Methods:** We performed a multi-level analysis of *HOXB13* gene expression using bulk RNA-sequencing data from GTEx and TCGA datasets. Expression profiles were evaluated across normal human tissues, multiple cancer types, and stratified by patient race. Differential expression in PRAD was assessed, and survival analysis was conducted using Kaplan–Meier curves based on *HOXB13* expression and Gleason scores. In parallel, functional variant annotation was performed using eight biological scoring criteria (seven in silico prediction tools plus GTEx tissue expression) to identify high-impact SNPs. **Results:** *HOXB13* exhibited tissue-specific expression, with the highest levels observed in normal prostate tissue. In cancer expression profiling, *HOXB13* was markedly overexpressed in PRAD tumor samples compared to normal prostate and other cancer types. This overexpression was consistent across races, with the highest median levels noted in African American patients. Heatmap analysis confirmed strong upregulation of *HOXB13* and other tumor-associated genes (e.g., *TAGLN*, *S100A6*) in PRAD, while some genes (e.g., *ARHGAP10*, *PITPNC1*) were downregulated. Kaplan–Meier survival analysis revealed that patients with high *HOXB13* expression had better survival outcomes, particularly when combined with higher Gleason scores—a paradoxical finding suggesting *HOXB13* may mark less aggressive tumor phenotypes or reflect differentiation status, as its expression reportedly decreases in advanced/neuroendocrine PRAD subtypes associated with poor prognosis. Finally, a scoring system based on 12 annotation tools identified 23 SNPs with a perfect score of 8, including rs1207429780 and rs748782183, indicating strong potential for pathogenicity and functional relevance. **Conclusion:** *HOXB13* is a prostate-specific gene with significant overexpression in PRAD and potential prognostic value. Its expression is consistent across racial groups and correlates with improved survival when highly expressed. The identification of high-scoring pathogenic SNPs further underscores its importance in PRAD biology. These findings support *HOXB13* as a promising biomarker and potential therapeutic target in PRAD.

Keywords: *HOXB13*, Expression, Prognosis, Pathogenic Variants, Prostate Adenocarcinoma

* Corresponding author. e-mail: lalu.irham@pharm.uad.ac.id.

1. Introduction

Prostate Adenocarcinoma (PRAD) ranks as the second most frequently diagnosed cancer and stands as the fifth leading cause of cancer-related mortality among men worldwide. According to the most recent Global Cancer Observatory (GLOBOCAN 2022) estimates, approximately 1.47 million new cases of prostate cancer were diagnosed globally in 2022, accompanied by nearly 397,000 deaths (Bray *et al.*, 2024; Shah *et al.*, 2025). This substantial global burden underscores PRAD significant impact on men's health across diverse populations and regions. The primary contributors to morbidity and mortality in individuals with PRAD are the long-term and diverse complications arising from the disease, with genomic factors playing a pivotal role. Genetic alterations, including inherited mutations and somatic genomic changes, significantly influence PRAD development, progression, and response to treatment. These genomic factors interact with environmental exposures and other determinants, collectively shaping the clinical outcomes and complications experienced by patients (Chu *et al.*, 2025). Understanding the genomic landscape of PRAD is, therefore, critical for improving prognostication and tailoring personalized therapeutic strategies (Pritchard *et al.*, 2016; Wang *et al.*, 2023).

Single-nucleotide polymorphisms (SNPs) contribute to the development and progression of PRAD by affecting gene expression, messenger RNA (mRNA) stability, and translational efficiency. These genetic variations can alter cellular functions and pathways critical to tumorigenesis. For example, genome-wide association studies have identified over 100 common SNPs linked to PRAD susceptibility, many of which reside in noncoding regions that regulate gene expression (Allemailem *et al.*, 2021). Specific SNPs, such as those in the 8q24 region, have been strongly associated with increased risk and may serve as valuable biomarkers for diagnosis and therapy. Moreover, SNPs influence key biological processes including steroid metabolism, oxidative stress, cell adhesion, angiogenesis, cell cycle regulation, and DNA repair, all of which are implicated in PRAD pathogenesis (Allemailem *et al.*, 2021). Functional studies have demonstrated that certain SNP variants can modulate biochemical activities relevant to tumor growth and metastasis, exemplified by the rs17632542 variant affecting PSA function and clinical outcomes (Srinivasan *et al.*, 2024). Additionally, interactions between SNPs in genes involved in angiogenesis, such as *MMP16*, *ROBO1*, *CSF1*, and *EGFR*, have been linked to aggressive PRAD phenotypes, highlighting the complex genetic architecture underlying disease progression (Lin *et al.*, 2013). Understanding how SNPs impact gene regulation and cellular pathways is essential for improving risk prediction, screening, and personalized management of PRAD. Among the genetic factors implicated in prostate cancer pathogenesis, *HOXB13*, a prostate-lineage transcription factor, has emerged as a key regulator of prostate development and tumorigenesis. *HOXB13* exhibits prostate-specific expression and plays critical roles in cellular differentiation, androgen signaling, and tumor progression (Patel *et al.*, 2024). Recent interactome analyses in prostate cancer cells identified multiple *HOXB13* protein partners (including *TBX3*) and confirmed its functional

relevance for prostate cancer cell growth (Erokhin *et al.*, 2025). Germline mutations in *HOXB13*, particularly the G84E variant, are strongly associated with hereditary prostate cancer predisposition across diverse populations, providing a compelling biological rationale for its systematic investigation in PRAD (Heise *et al.*, 2019). This study focuses specifically on *HOXB13* to comprehensively characterize its expression patterns, prognostic significance, and pathogenic variants across ethnically diverse PRAD cohorts.

The identification of genomic variants associated with cancer has traditionally concentrated on mutations within protein-coding genes, particularly somatic mutations that drive tumor development. Databases such as the Catalogue of Somatic Mutations In Cancer (COSMIC) have been instrumental in prioritizing these somatic variants within coding regions for further study and clinical relevance (Tate *et al.*, 2019). However, it is now recognized that the vast majority of the human genome consists of non-coding regions, which include non-coding RNAs and other regulatory elements that are not yet fully annotated. Importantly, over 90% of genomic variants are located within these non-coding regions (Maurano *et al.*, 2012). This realization highlights the critical need to investigate the functional impact of genomic variants outside of protein-coding sequences, especially germline variants that may predispose individuals to cancer. Understanding how these non-coding variants contribute to cancer susceptibility and progression is essential for comprehensive genomic profiling and the development of personalized medicine approaches. Emerging research has revealed that germline variants—heritable genetic alterations present in normal cells—play a significant role not only in increasing cancer susceptibility but also in influencing tumor progression and patient outcomes across various cancer types, including PRAD. These inherited variants can interact with somatic mutations, shaping the landscape of genomic aberrations that drive oncogenesis. For instance, germline mutations in tumor suppressor genes such as *BRCA1* and *BRCA2* have been shown to predispose individuals to cancer and are often accompanied by loss of the wild-type allele in tumors, consistent with Knudson's two-hit hypothesis (Chatrath *et al.*, 2021; Lichtenstein *et al.*, 2000; Huang *et al.*, 2018).

Genetic variants play a crucial role in the development of cancer and other disease by inducing alterations at the molecular level (Halla and Wathiq, 2025; Usha and Rajeev, 2025; Sözen, 2024). These changes can manifest as single nucleotide mutations or larger chromosomal abnormalities. While the majority of these mutations, known as passenger mutations, accumulate during cancer progression without directly contributing to tumorigenesis, a smaller subset—referred to as driver mutations—actively promotes oncogenesis by conferring growth advantages to cancer cells (Vogelstein *et al.*, 2013). Distinguishing these driver mutations from the extensive background of passenger mutations is essential for understanding cancer biology and identifying therapeutic targets. In this study, we employed bioinformatics approaches to systematically identify putative genomic variants that are associated with increased susceptibility to PRAD. This study addresses these gaps through a comprehensive, multi-level analysis of *HOXB13* expression, prognosis, and pathogenic variants in PRAD across

ethnically diverse populations. We hypothesized that *HOXB13* demonstrates prostate-specific overexpression with complex prognostic significance, systematically characterizing: (1) tissue/cancer expression profiles; (2) race-stratified survival associations with Gleason scores; and (3) high-confidence pathogenic SNPs via eight-criteria bioinformatics pipeline.

2. Method

2.1. Identification and Functional Annotation of *HOXB13*

Despite advances in genomic technologies, interpreting the functional significance of genomic variants, particularly those with potential pathogenic effects, remains a complex challenge. To address this, we developed an integrative bioinformatics pipeline to identify and prioritize potentially pathogenic variants in the *HOXB13* gene, which has been implicated in PRAD.

Variant data for *HOXB13* were retrieved from the ENSEMBL database (accessed: Oct 15, 2025), and analysis was limited to non-synonymous SNPs, especially missense variants, due to their potential to alter protein function. Disease-related annotations were further obtained from PubMed, the Genome-Wide Association Studies (GWAS) catalog (accessed: Oct 18, 2025) (Buniello *et al.*, 2019), and the Phenome-Wide Association Studies (PheWAS) catalog (accessed: Oct 19, 2025) to evaluate prior evidence of clinical or phenotypic associations (Denny *et al.*, 2013).

To functionally assess the impact of variants, we incorporated several widely used in silico prediction tools: (I) Sorting Intolerant From Tolerant (SIFT) predicts whether an amino acid substitution affects protein function based on sequence homology and amino acid properties (Ng and Henikoff, 2003); (II) Polymorphism Phenotyping (PolyPhen) assesses the likely structural and functional consequences of substitutions using physical and evolutionary information (Adzhubei *et al.*, 2010); (III) Combined Annotation Dependent Depletion (CADD) integrates multiple genomic annotations into a composite score estimating the deleteriousness of single nucleotide variants (SNVs) and insertion/deletion (indels) mutations (Kircher *et al.*, 2014); (IV) Rare Exome Variant Ensemble Learner (REVEL) combines predictions from various algorithms to improve the accuracy of missense variant classification (Ioannidis *et al.*, 2016); (V) META-LR uses logistic regression to aggregate outputs from multiple prediction methods for more reliable pathogenicity estimates (Ioannidis *et al.*, 2016; Dong *et al.*, 2015); (VI) Mutation Assessor evaluates functional impact based on evolutionary conservation across species (Reva *et al.*, 2011); and (VII) AlphaMissense applies deep learning models and AlphaFold protein structure predictions to infer how missense variants may influence protein structure and function (Frazer *et al.*, 2021). Additionally, tissue-specific expression was evaluated using the Genotype-Tissue Expression (GTEx) Portal (GTEx Consortium, accessed: Oct 20, 2025), which confirmed that *HOXB13* exhibits its highest expression levels in prostate tissue. This expression profile provided functional validation of gene relevance to prostate-specific biology and was incorporated as a supporting criterion in variant prioritization.

Each variant was scored using a binary system (1 = deleterious, 0 = benign) across eight biological scoring criteria comprising seven in silico prediction tools (SIFT, PolyPhen, CADD, REVEL, META-LR, Mutation Assessor, AlphaMissense) plus GTEx prostate tissue expression. The resulting cumulative score (range: 0-8) classified variant pathogenicity, with scores of 8 identifying high-confidence pathogenic candidates. This integrated scoring approach enabled systematic prioritization based on sequence conservation, biochemical effect, structural integrity, disease association, and tissue-specific expression. A schematic overview of this analytical workflow is presented in Figure 1.

2.2. Gene Expression Analysis of *HOXB13*

To evaluate the normal tissue expression of *HOXB13*, RNA-sequencing data were obtained from the GTEx Portal. Transcript abundance was assessed across multiple human tissues to determine baseline expression profiles. This analysis confirmed that *HOXB13* is highly expressed in prostate tissue, reinforcing its potential role in prostate-related pathophysiology. To further examine gene expression in the disease context, we used the UALCAN web resource to access the Cancer Genome Atlas (TCGA) PRAD dataset (accessed: Oct 22, 2025) (Chandrashekar *et al.*, 2022). *HOXB13* expression levels were compared between the tumor and adjacent normal prostate tissues. This comparison enabled the evaluation of differential expression and the identification of potential dysregulation of *HOXB13* in malignancy. Expression patterns were visualized across sample groups to highlight tumor-specific trends.

2.3. Prognostic Evaluation based on *HOXB13* Expression

To assess the prognostic value of *HOXB13*, gene expression data from TCGA-PRAD were further analyzed in relation to clinical outcomes. Using UALCAN's clinical analysis tools (accessed: Oct 25, 2025) (Chandrashekar *et al.*, 2022), we correlated *HOXB13* expression levels with Gleason scores and overall survival. Patients were stratified by high versus low/medium *HOXB13* expression and further categorized based on Gleason grade (ranging from 6 to 10). Kaplan-Meier survival curves were generated to evaluate differences in survival outcomes, with statistical significance determined using log-rank tests. In addition, expression levels of *HOXB13* were compared between racial subgroups, specifically Caucasian and African American prostate cancer patients, to explore potential population-level disparities. Although limited by a small sample size in the African American cohort, this analysis served as a preliminary assessment of race-based expression variation in *HOXB13* within PRAD.

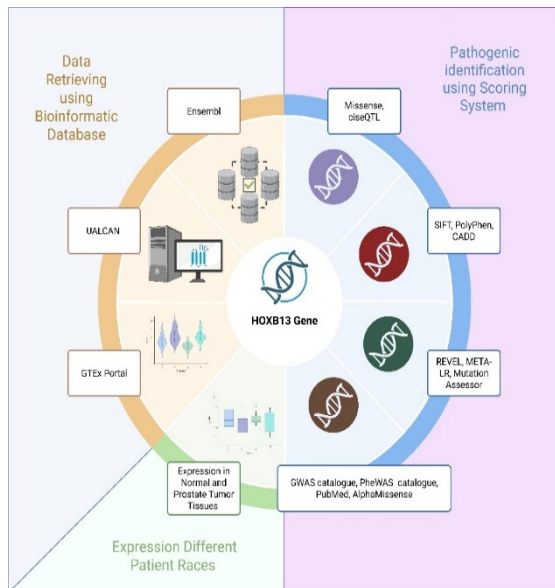


Figure 1. Study workflow and bioinformatics analysis pipeline. The workflow of the present study illustrates the sequential integration of multiple publicly available databases and bioinformatics prediction tools. These included ENSEMBL, PubMed, the Genome-Wide Association Studies (GWAS) Catalog, and the Phenome-Wide Association Studies (PheWAS) Catalog, along with several established variant-effect prediction tools, including SIFT, PolyPhen, CADD, REVEL, META-LR, Mutation Assessor, and AlphaMissense.

3. Result

3.1. Gene expression of *HOXB13* revealed highest expressed on the prostate tissue

Gene expression analysis of *HOXB13* across various human tissues reveals that it is most highly expressed in the prostate (Figure 2). This is clearly illustrated by the significantly elevated expression levels in the prostate compared to all other tissues examined. Moderate expression is also observed in related reproductive and urinary tissues such as the vagina, cervix, bladder, and testis, but these levels are substantially lower than in the prostate. In contrast, most other tissues—including liver, lung, heart, brain, muscle, and blood—show minimal to negligible expression of *HOXB13*. This distinct tissue-specific expression pattern underscores the critical role of *HOXB13* in prostate biology, where it is involved in regulating development, differentiation, and androgen receptor signalling. Given its predominant expression in the prostate and its known association with PRAD risk, *HOXB13* represents an important biomarker and potential therapeutic target. Furthermore, its limited expression in non-prostate tissues suggests that interventions targeting *HOXB13* may have minimal off-target effects, enhancing their clinical appeal. Overall, these findings highlight the specialized function of *HOXB13* in the prostate and its significance in PRAD research and treatment strategies.

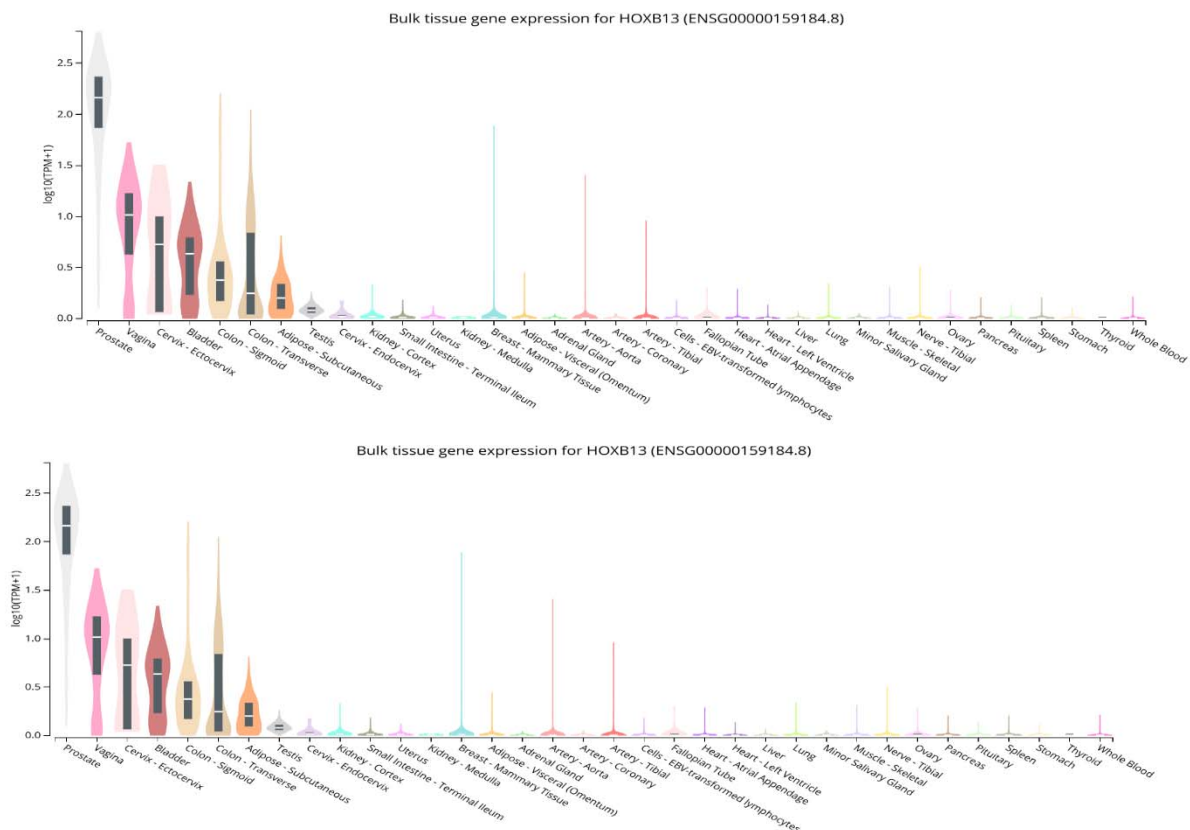


Figure 2. Violin plot depicting the bulk tissue gene expression levels of *HOXB13* gene across multiple human tissues. The y-axis represents the log₁₀-transformed TPM (Transcripts Per Million) values plus one [$\log_{10}(\text{TPM}+1)$], providing normalized expression data. Each violin illustrates the distribution of expression values within a given tissue, with the white dot indicating the median expression and the black bar representing the interquartile range. The plot reveals that *HOXB13* is most highly expressed in prostate tissue compared to all other tissues, highlighting its tissue-specific expression pattern. Moderate expression is observed in select reproductive and urinary tissues, while expres-

sion in other tissues remains low or negligible. This expression profile underscores the gene's specialized role in prostate biology and its relevance to PRAD research.

3.2. Expression profile of *HOXB13* across human cancers

To evaluate the tissue- and cancer-specific expression patterns of *HOXB13*, we analyzed transcriptomic data from The TCGA across 33 tumor types. Gene expression levels were normalized and presented as $\log_2(\text{TPM} + 1)$ for comparative analysis.

Figure 3A illustrates *HOXB13* expression in matched tumors (red boxes) and normal tissue samples (blue boxes) across various cancer types. Among all tumor types examined, PRAD exhibited the highest *HOXB13* expression. PRAD tumor tissues demonstrated markedly elevated median expression levels compared to their normal counterparts, as well as to all other tumor types. This strong differential underscores the prostate-specific and cancer-enriched expression of *HOXB13*, supporting its role as a lineage-survival gene and a candidate biomarker for PRAD.

In contrast, for most other cancer types—including bladder (BLCA), breast (BRCA), colon (COAD), lung (LUAD, LUSC), liver (LIHC), and kidney cancers (KICH, KIRC, KIRP)—*HOXB13* expression remains low in both tumor and normal samples, as indicated by the short height of the boxes and low median values. Notably, in some cancers such as colon adenocarcinoma (COAD) and rectum adenocarcinoma (READ), normal tissues show slightly higher *HOXB13* expression compared to their corresponding tumors, but these levels are still much lower than those observed in PRAD. A few other cancer types, such as sarcoma (SARC) and skin cutaneous melanoma (SKCM), show moderate expression in tumor samples, but

again, these do not approach the levels seen in PRAD. The consistently low expression outside of prostate tissues reinforces the gene's tissue-specific regulatory function.

Figure 3B presents *HOXB13* expression exclusively across TCGA tumor samples, allowing a focused comparison of gene activity within the malignant tissue landscape. As expected, PRAD remained the dominant outlier, with uniformly high expression across nearly all tumor samples. This finding aligns with the known biological function of *HOXB13* in prostate development and androgen receptor signaling. Several other tumor types, including rectum adenocarcinoma (READ), pheochromocytoma and paraganglioma (PCPG), uveal melanoma (UVM), and uterine corpus endometrial carcinoma (UCEC), showed moderate expression with wider inter-sample variability, suggesting potential subtype-specific roles or regulatory heterogeneity. Conversely, cancers such as glioblastoma (GBM), lower-grade glioma (LGG), kidney chromophobe (KICH), diffuse large B-cell lymphoma (DLBC), and thymoma (THYM) show very low or near-absent *HOXB13* expression, suggesting the gene has limited or no involvement in these tumor types. Some cancers, like breast cancer (BRCA), bladder cancer (BLCA), and colon adenocarcinoma (COAD) demonstrate moderate expression with wider variation among samples, which may reflect tumor heterogeneity or subtype differences. In summary, this pattern suggests that *HOXB13* is a gene with selective overexpression, most notably in PRAD, and may hold potential as a biomarker or therapeutic target in certain malignancies.

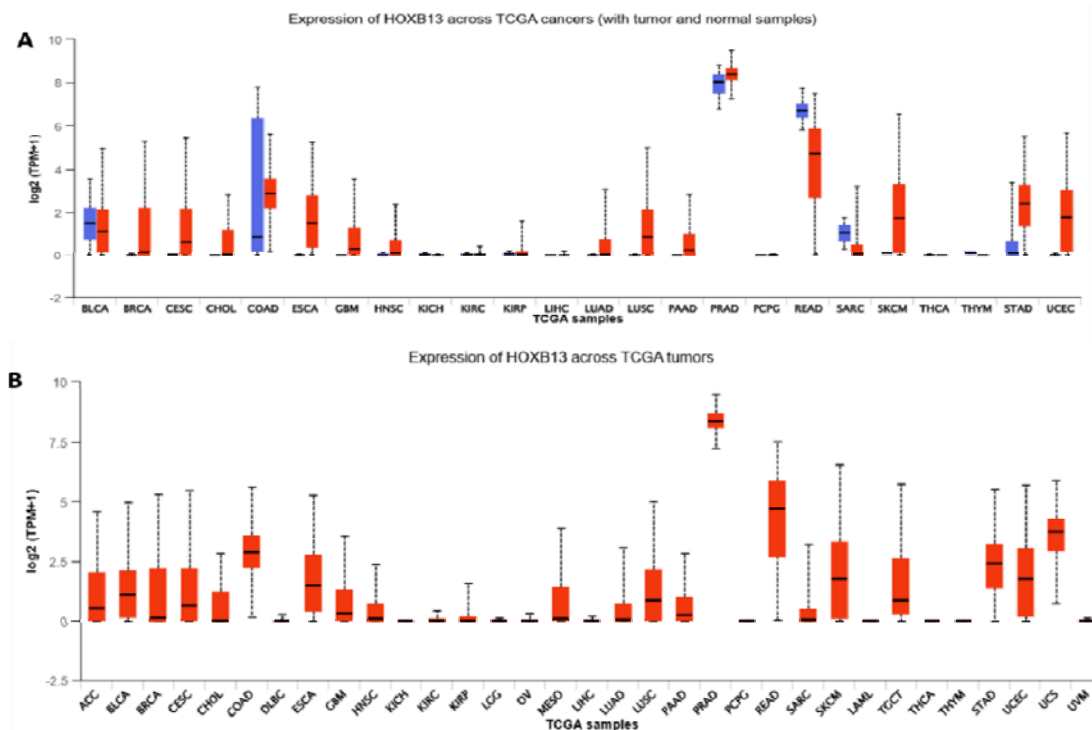


Figure 3. A. Expression of the *HOXB13* gene across TCGA cancers with tumor and normal samples. Box plots represent *HOXB13* gene expression levels measured as $\log_2(\text{TPM} + 1)$ across multiple cancer types from TCGA. Red boxes indicate tumor samples, while blue boxes represent normal tissue samples. Cancer types are labeled along the x-axis using TCGA abbreviation codes (e.g., BLCA: Bladder Cancer, PRAD: Prostate Adenocarcinoma). The plot demonstrates that *HOXB13* is most highly expressed in PRAD compared to other cancer types, with notable expression also observed in rectum adenocarcinoma (READ) and cholangiocarcinoma (CHOL) normal tissues. Each box represents the interquartile range, with the horizontal line indicating the median expression value, and whiskers extending to show data distribution.

tion within 1.5 times the interquartile range. **B.** Expression of *HOXB13* across various TCGA tumor types. Gene expression levels are shown as $\log_2(\text{TPM}+1)$. *HOXB13* is highly expressed in PRAD, with variable expression in other cancers and minimal expression in several tumor types.

3.3. Analysis of *HOXB13* Expression in Normal and Prostate Tumor Tissues Across Different Patient Races

Figure 4A presents a comparison of *HOXB13* gene expression between normal prostate tissues and primary prostate tumor tissues (PRAD) based on TCGA data. Expression levels are measured in Transcripts Per Million (TPM). The results show that *HOXB13* expression is significantly higher in primary tumor samples (n=497) compared to normal tissues (n=52). The median expression in tumor samples is notably elevated, with a narrower distribution range, indicating consistent overexpression across most cancer cases. In contrast, normal tissues exhibit a broader range of expression with several samples showing low or near-zero levels. This suggests that *HOXB13* is markedly upregulated in PRAD and may play a role in tumor development or progression. These findings support the potential use of *HOXB13* as a diagnostic biomarker or therapeutic target in PRAD.

Figure 4B displays *HOXB13* expression in PRAD tumor samples stratified by patient race, again based on

TCGA data. Samples are categorized into three groups: normal tissue (n = 52), Caucasian patients (n = 147), and African American patients (n = 6). The y-axis shows gene expression in TPM. The analysis reveals elevated *HOXB13* expression in both racial groups compared to normal tissues. Median expression is highest in African American patients, followed closely by Caucasian patients, while the normal group shows the lowest expression overall. Interquartile ranges and whiskers indicate variability within each group, with tumor samples displaying greater expression levels and variability than normal samples. Although the African-American group is represented by a small sample size (n = 6), the trend suggests that *HOXB13* overexpression in PRAD may occur across racial groups, potentially with slightly higher expression in African-American patients. These findings warrant further validation with larger sample sizes but highlight the consistent upregulation of *HOXB13* in PRAD and its relevance across different racial backgrounds.

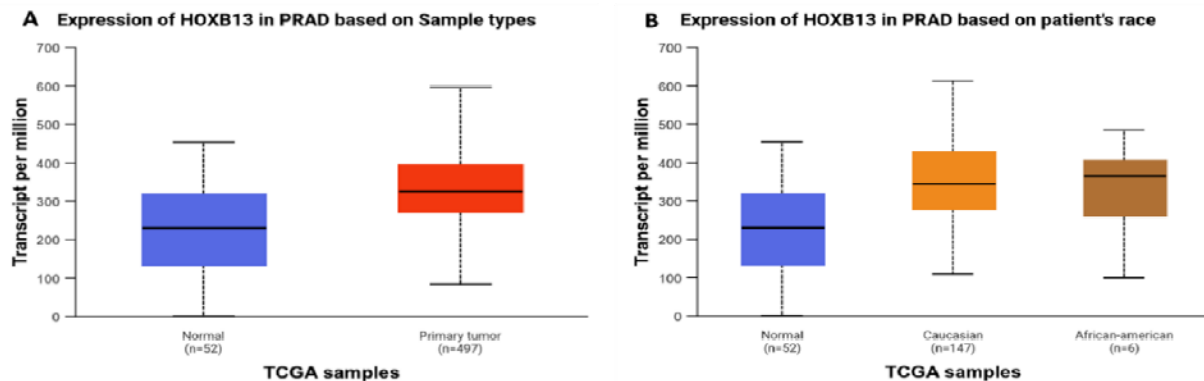


Figure 4. A. Box plot showing *HOXB13* expression (TPM) in normal (n=52) and primary prostate tumor tissues (n=497) from the TCGA PRAD dataset. *HOXB13* is significantly overexpressed in tumors, suggesting a role in PRAD. B. *HOXB13* expression in normal tissue and PRAD tumors from Caucasian (n=147) and African-American (n=6) patients. Elevated expression is observed in both tumor groups, with the highest median in African-American patients. Data accessed on 10 June 2025 from UALCAN.

3.4. Heatmap of Gene Expression Patterns in Normal and Tumor Samples of PRAD

Figure 5A presents a heatmap showing the expression pattern of selected input genes in PRAD, comparing normal and tumor samples from the TCGA dataset. The x-axis represents individual samples categorized into normal and tumor groups, while the y-axis lists the input genes, including *HOXB13*, *SCCPDH*, *TRNT1*, *GDPPD1*, and others. Gene expression levels are visualized using a color scale based on $\log_2(\text{TPM}+1)$, where blue indicates low expression and red indicates high expression. Based on the heatmap, it is evident that *HOXB13* is highly overexpressed in tumor tissues compared to normal tissues, as shown by the strong red signal in the tumor group and minimal expression (blue) in the normal group. This reinforces the earlier finding that *HOXB13* is significantly upregulated in PRAD. Several other genes also show altered expression between normal and tumor samples. For example, *GDPPD1* and *C15orf37* exhibit relatively higher expression in normal samples (deep blue in tumors), suggesting possible downregulation in cancer. In contrast, some genes like *UBE2Z* and *TRPC4AP* show more uniform or

slightly increased expression in tumors. Overall, this heatmap provides a comparative overview of gene expression changes in PRAD, highlighting potential candidate genes, especially *HOXB13*, that may contribute to tumor development or progression. These patterns can inform further investigations into the biological functions and clinical relevance of these genes in PRAD.

Figure 5B further highlights genes with marked differential expression. In addition to *HOXB13*, genes such as *TAGLN*, *S100A6*, and *EMILIN1* are consistently upregulated in tumor tissues, indicated by strong red coloration. These genes may be involved in oncogenic signaling, cytoskeletal remodeling, or stress response pathways relevant to prostate cancer progression. In contrast, genes including *NAAALDL1*, *PITPNC1*, and *ARHGAP10* exhibit reduced expression in tumor samples, suggesting downregulation during malignant transformation and potential roles in maintaining normal cellular homeostasis. Overall, the heatmap analyses in both panels provide a comprehensive view of transcriptional changes between normal and tumor prostate tissues. The observed expression shifts emphasize the deregulation of genes involved in structural organiza-

tion, signaling, and tumor biology. Notably, the prominent upregulation of *HOXB13*, along with several co-regulated genes, highlights a panel of candidate biomarkers and

therapeutic targets that warrant further functional validation in PRAD research.

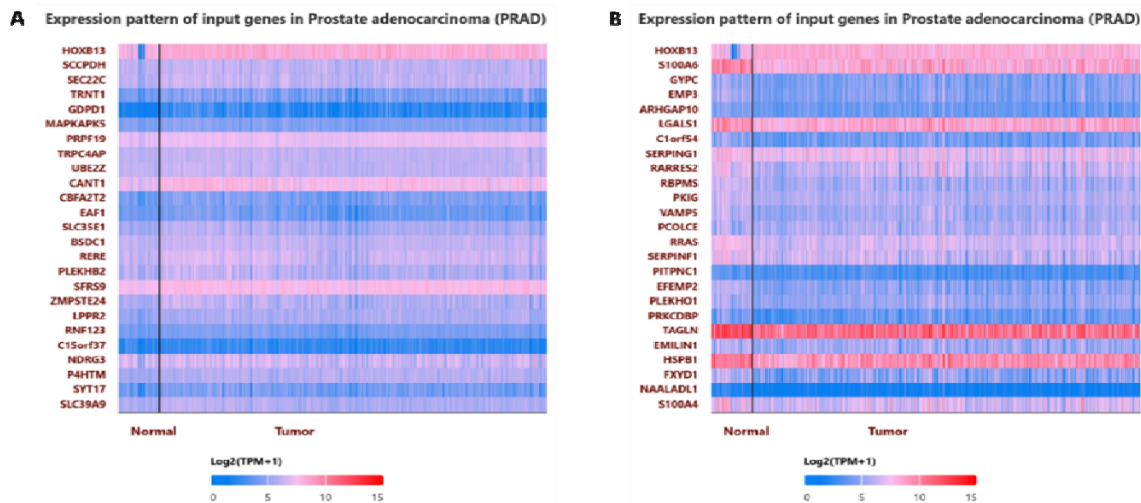


Figure 5. Heatmaps displaying the expression patterns of selected genes in normal and tumor prostate tissues from the TCGA PRAD dataset. Panel A shows gene expression as $\log_2(\text{TPM}+1)$ with each column representing a sample and each row a gene. *HOXB13* is notably overexpressed in tumor samples compared to normal tissue, suggesting its role in Prostate Adenocarcinoma. Panel B highlights additional genes with differential expression: *HOXB13*, *TAGLN*, *S100A6*, and *EMILIN1* are upregulated in tumors, whereas *PITPNC1* and *NAAALDL1* are downregulated. These expression patterns identify potential biomarkers and therapeutic targets involved in PRAD progression.

3.5. Survival Analysis of PRAD Patients Based on *HOXB13* Expression and Gleason Score

Figure 6 presents a Kaplan–Meier survival analysis illustrating the combined effect of *HOXB13* gene expression levels and Gleason scores on overall survival in PRAD patients. The x-axis represents time in days, while the y-axis shows survival probability. The patients are stratified into subgroups based on *HOXB13* expression (high vs. low/medium) and Gleason score (ranging from 6 to 10). Each curve represents a different subgroup, and the number of patients (n) in each group is indicated in the legend. A statistically significant p-value of 0.00031 indicates a meaningful difference in survival among the groups. From the plot, it is clear that patients with low or medium *HOXB13* expression and high Gleason scores (particularly 9 and 10) have poorer survival outcomes, as shown by steeper declines in survival probability. For instance, the "Low/Medium expression + Gleason-score10" group (n=4)

and "Low/Medium expression + Gleason-score9" group (n=98) show the most rapid decrease in survival, suggesting a negative prognostic impact of low *HOXB13* expression when combined with high tumor aggressiveness. In contrast, patients with high *HOXB13* expression, even with higher Gleason scores, tend to have better survival probabilities, suggesting a potential protective or prognostic role of elevated *HOXB13* expression in certain contexts.

Overall, current findings highlighted, this analysis underscores the potential of *HOXB13* expression level as a prognostic biomarker in PRAD, especially when interpreted in conjunction with the Gleason score. The findings suggest that patients with high *HOXB13* expression tend to have better survival outcomes, even when their tumors are histologically more aggressive, highlighting the clinical relevance of integrating molecular and pathological data in patient risk stratification.

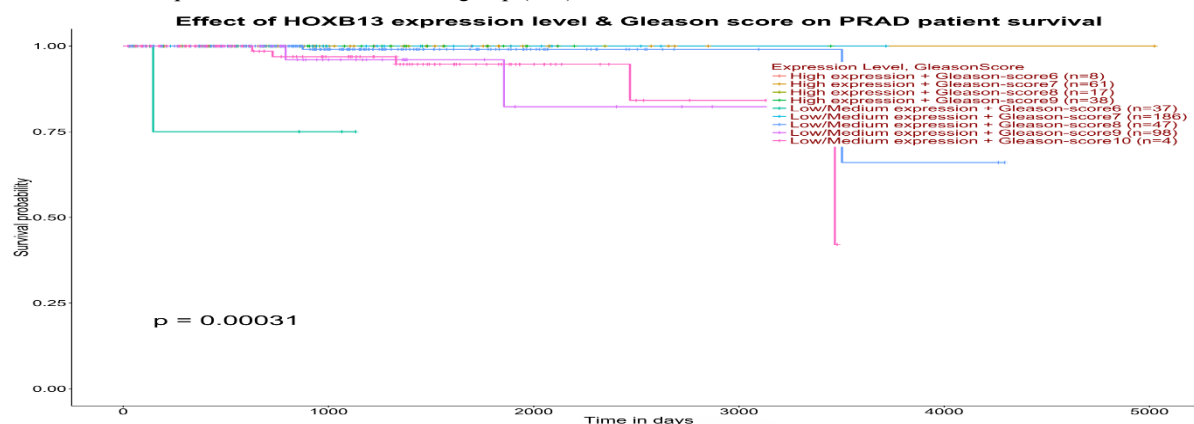


Figure 6. Kaplan–Meier plot showing overall survival of PRAD patients stratified by *HOXB13* expression levels (high vs. low/medium) and Gleason scores (6–10) using TCGA data. Each colored line represents a patient subgroup, with sample size (n) indicated in the legend. Patients with low/medium *HOXB13* expression and high Gleason scores (especially 9 and 10) exhibit significantly poorer survival compared to other groups. In contrast, patients with high *HOXB13* expression generally show better survival outcomes. The log-rank test yields a significant

cant p-value of 0.00031, indicating a statistically meaningful difference in survival across subgroups. These results suggest that *HOXB13* expression may serve as a potential prognostic biomarker in PRAD when combined with Gleason grading.

3.6. Pathogenic variant identification using a functional scoring system

Figure 7 presents a bar chart summarizing the number of *HOXB13* SNPs identified using various biological and computational criteria, reflecting their functional and clinical relevance. The x-axis denotes different tools and databases used for SNP evaluation, while the y-axis represents the number of SNPs identified under each scoring criterion. Among all tools, AlphaMissense identified the largest number of SNPs (843), followed closely by the Missense annotation category (835), emphasizing the focus on coding variants predicted to alter amino acid sequences. Established predictive tools such as PolyPhen and SIFT also yielded a high number of variants (592 and 526, respectively), highlighting their ongoing utility in assessing the potential functional consequences of amino acid substitutions.

Ensemble and integrative tools, including MetaLR (467 SNPs), REVEL (334 SNPs), and CADD (101 SNPs), further contribute to the identification of functionally relevant variants, demonstrating the value of combining multiple predictive approaches. In contrast, tools with limited variant yield included Mutation Assessor (34 SNPs), PubMed (5 SNPs), GWAS (1 SNP), and PheWAS (0 SNPs). The relatively low representation from clinical association databases and literature-based resources reflects the current gap between computational prediction and experimentally validated or population-level data.

The data highlights a computational prioritization framework in which algorithm-based predictions, particularly those from AlphaMissense, play a central role in fil-

tering genetic variants. Although these tools are highly effective at identifying variants with potential functional effects, the limited number of SNPs supported by clinical or literature evidence points to several important needs: expanded experimental validation studies, greater representation of diverse populations in genetic association databases, and the incorporation of multi-omics data beyond purely computational predictions. While this multi-faceted strategy efficiently narrows down candidate variants, it also reveals significant gaps between in silico predictions and variants with confirmed clinical relevance.

The result highlights a group of SNPs that received the highest functional annotation score of 8, indicating their strong potential for biological impact. These SNPs fulfilled all eight evaluation criteria used in the scoring system, namely Missense, SIFT, PolyPhen, CADD, REVEL, META-LR, Mutation Assessor, and AlphaMissense. The SNPs that achieved this top score include: rs1207429780, rs1365586165, rs1371765250, rs155558450, rs1597932515, rs1597932533, rs1597932564, rs1597932808, rs1597932884, rs2038215417, rs2143065331, rs2143066906, rs2143067303, rs747483887, rs748782183, rs749518336, rs754598240, rs757776260, rs778197145, and rs2143065697. Notably, some SNPs such as rs748782183, rs778197145, and rs1597932533 appear more than once in the dataset, suggesting potential occurrences of multiple allelic combinations or repeated annotations. The consistently high scores across multiple predictive tools emphasize the potential significance of these variants in genetic or disease association studies, warranting further functional validation.

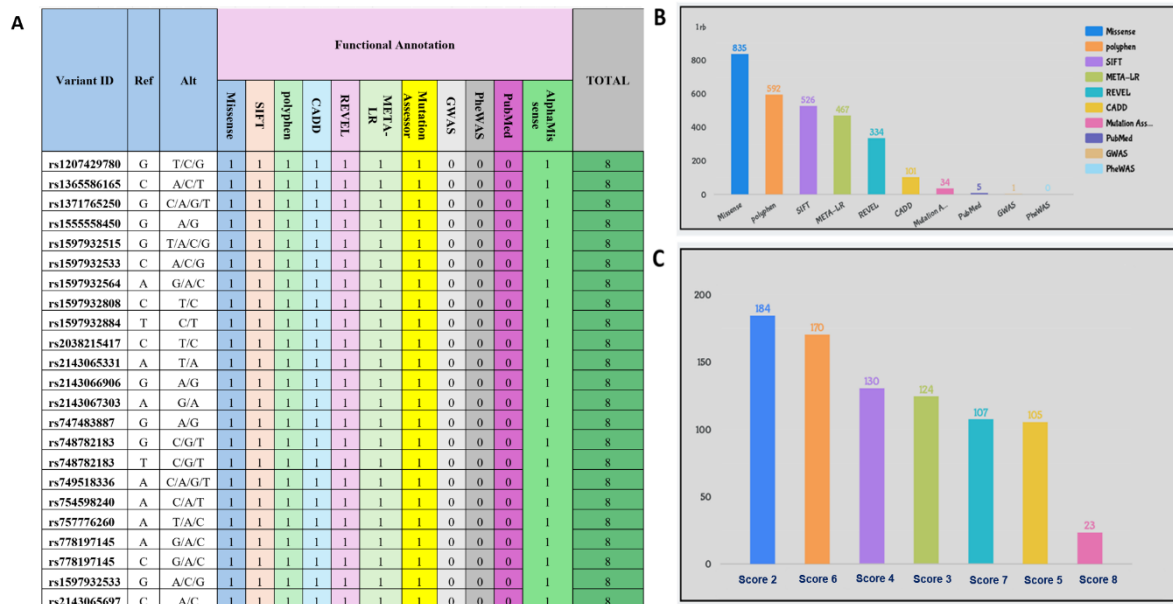


Figure 7. A-B. Bar chart showing the number of Single Nucleotide Polymorphisms (SNPs) identified under various biological scoring criteria used to evaluate their functional significance. The criteria include computational prediction tools (e.g., SIFT, PolyPhen, CADD, REVEL, Mutation Assessor, MetaLR), clinical and literature-based sources (e.g., GWAS, PubMed, PheWAS), and integrated models such as AlphaMissense. AlphaMissense and Missense categories contribute the highest number of SNPs (843 and 835, respectively), highlighting their prominent roles in functional annotation. In contrast, categories like PheWAS and GWAS report minimal or no SNPs, indicating limited coverage or association. This scoring distribution reflects a comprehensive approach to SNP prioritization across multiple biological relevance indicators. C. List of high-confidence SNPs with maximum functional impact scores. The figure displays SNPs that achieved the highest possible score (8 out of 8) across multiple biological annotation tools, including Missense, SIFT, PolyPhen, CADD, REVEL, ME-

TA-LR, Mutation Assessor, and AlphaMissense. These variants are prioritized as likely to have significant functional consequences and are potential candidates for further investigation in disease-related studies.

4. Discussion

This study provides a comprehensive multi-level evaluation of *HOXB13* expression across normal human tissues, multiple cancer types, patient racial backgrounds, and its potential clinical relevance in prostate adenocarcinoma (PRAD). Consistent across GTEx and TCGA datasets, *HOXB13* exhibits a striking prostate-restricted expression pattern, with the highest expression observed in normal prostate tissue and marked upregulation in PRAD tumors. The near absence of *HOXB13* expression in non-prostatic tissues and non-PRAD malignancies reinforces its tissue specificity and establishes *HOXB13* as a robust diagnostic biomarker for confirming prostatic origin, particularly in metastatic tumors of unknown primary.

Our analysis of *HOXB13* expression across cancers using TCGA data confirmed its marked upregulation in PRAD tumor samples compared to normal prostate tissue and other tumor types, consistent with previous studies demonstrating *HOXB13*'s prostate-specific expression and its role in PRAD progression (Patel *et al.*, 2024; Dupont *et al.*, 2021; Zabalza *et al.*, 2015). The negligible expression of *HOXB13* in non-PRAD further supports its tissue-restricted nature and validates its use as a diagnostic biomarker for PRAD, particularly in confirming prostatic origin in metastatic tumors (Patel *et al.*, 2024). Our race-stratified analysis showed elevated *HOXB13* expression in both Caucasian and African-American PRAD patients, with a trend toward higher median expression in African-American men, aligning with reports highlighting *HOXB13*'s universal role across ethnic groups despite some limitations due to smaller sample sizes (Patel *et al.*, 2024). Kaplan–Meier survival analysis revealed that high *HOXB13* expression correlates with improved overall survival, particularly in patients with high Gleason scores, suggesting a complex tumor differentiation role. This paradoxical finding aligns with reports that *HOXB13* expression decreases in advanced/neuroendocrine PRAD subtypes associated with poor prognosis (Kanayama *et al.*, 2024). High *HOXB13* thus identifies Gleason-aggressive but High *HOXB13* expression may therefore identify tumors with high Gleason scores but a more differentiated molecular phenotype, positioning it as a critical complement to histopathological grading (Patel *et al.*, 2024). Conversely, patients with low or medium *HOXB13* expression and high Gleason scores had poorer outcomes, underscoring *HOXB13*'s potential as a dual prognostic marker when combined with histopathological data (Dupont *et al.*, 2021; Zabalza *et al.*, 2015). Collectively, these results reinforce *HOXB13*'s significance as both a diagnostic and prognostic factor in PRAD and highlight the need for further research into its mechanistic role across diverse populations.

Our heatmap analysis revealed significant upregulation of *HOXB13*, along with genes such as *TAGLN* and *S100A6*, in PRAD tumor tissues compared to normal samples, while genes like *ARHGAP10* and *PITPNC1* were downregulated. These findings align with prior studies that have reported *HOXB13*'s involvement in PRAD progression and its co-expression with genes implicated in cytoskeletal organization and tumor invasion (Egevad *et al.*,

2002). For instance, *S100A6* has been recognized for its role in tumor proliferation and metastasis, and *TAGLN* is known to influence cellular motility, both contributing to PRAD aggressiveness (Larkin *et al.*, 2012; Ohuchida *et al.*, 2005). The downregulation of *ARHGAP10* and *PITPNC1* further suggests disruption of signalling pathways critical for maintaining normal prostate epithelial architecture, consistent with observations in other genomic studies of PRAD (Gong *et al.*, 2019). Complementing the expression data, our comprehensive SNP functional annotation identified 23 variants with maximal scores across 12 predictive tools, including rs1207429780 and rs748782183. The convergence of predictions from tools such as AlphaMissense, SIFT, PolyPhen, and CADD supports the potential pathogenicity of these variants. This multi-tool consensus approach has been previously advocated as an effective strategy to prioritize SNPs for functional validation and risk association studies in cancer genomics (Ioannidis *et al.*, 2016; Ng and Henikoff, 2001). While some of these variants have yet to be characterized experimentally, their identification provides promising targets for future research aimed at elucidating the genetic underpinnings of PRAD susceptibility and progression.

HOXB13 is predominantly expressed in prostate tissue, playing a crucial role in prostate development and cancer risk, with limited expression in other tissues. The rare missense variant rs138213197 (G84E) in *HOXB13* is strongly associated with increased PRAD susceptibility, though it is underrepresented in large-scale GWAS and PheWAS databases due to its rarity (Bıyıkoglu *et al.*, 2024; Zhang *et al.*, 2016). Discovered through exome sequencing in primarily Caucasian populations, rs138213197 confers a significantly elevated PRAD risk, with multiple studies confirming odds ratios ranging from 3.3 to over 20 across diverse cohorts (Kote-Jarai *et al.*, 2015). While its association with breast cancer remains uncertain, the variant appears to be a founder mutation originating in Northern Europe around the 1800s, particularly prevalent in Nordic populations. This highlights the importance of rs138213197 in PRAD genetics and the need for focused studies beyond broad population databases (Witte *et al.*, 2013).

These convergent findings position *HOXB13* as a multifaceted biomarker with diagnostic specificity, prognostic utility, and therapeutic relevance in PRAD. The identification of high-impact germline variants further underscores its role in hereditary susceptibility, exemplified by the well-characterized G84E founder mutation (rs138213197). Future research should prioritize experimental validation of prioritized SNPs, single-cell resolution of tumor heterogeneity, and prospective clinical studies integrating *HOXB13* expression with histopathological and genomic profiling. Taken together, this multi-layered analysis reinforces the role of *HOXB13* as a prostate-specific oncogene with diagnostic, prognostic, and therapeutic relevance. Its selective overexpression in prostate tissue, correlation with clinical outcomes, and enrichment of high-impact pathogenic variants highlight its central role in PRAD biology. Nonetheless, further studies are needed to experimentally validate these candidate variants, delineate the mechanisms by which *HOXB13* contributes to tumor progression,

and assess its utility across diverse genetic and clinical contexts.

Strengths of this study include the use of comprehensive bulk RNA sequencing data from large publicly available databases (GTEx and TCGA), race-stratified expression profiling, integration of survival analysis with Gleason scoring, and a multi-algorithm SNP prioritization pipeline. However, several limitations should be acknowledged: the use of bulk RNA-seq data does not capture cell-type-specific expression or tumor heterogeneity; the observational nature of the analysis limits causal inference; and the small representation of certain racial and ethnic groups constrains generalizability. Additionally, while multiple computational tools were used for variant annotation, experimental validation remains essential to confirm biological impact as suggested by a previous study (Admanegara *et al.*, 2025).

5. Conclusion

This study demonstrates that HOXB13 is a prostate-restricted gene with consistent overexpression in PRAD across large transcriptomic datasets and observable associations with clinical outcomes. The identification of high-confidence pathogenic variants further supports its biological relevance in prostate cancer susceptibility and progression. Collectively, these findings suggest that HOXB13 represents a potential diagnostic and prognostic biomarker—and a candidate therapeutic target—in PRAD. However, clinical applicability will require further validation through functional experiments, single-cell analyses, and well-powered, multi-ethnic clinical studies. Continued investigation is, therefore, warranted to fully elucidate the mechanistic and translational significance of HOXB13 in prostate cancer.

The Acknowledgements

None

External funding

The authors gratefully acknowledge the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, for providing financial support for this research through grants No. 0209/DST/C3/DT.05.00/2026; No. 132/C3/DT.05.00/PL-Multitahun Lanjutan/2026; and No. 0348.07/LL5-INT/AL.04.03/2026.

Data availability statement

Data will be made available upon request.

Declaration of competing interests

The authors disclose no conflicting interests.

Ethics statement

This study did not require ethical approval as it did not involve human participants, animal subjects, or sensitive data that necessitated ethical review.

Author contribution statement

Lalu Muhammad Irham and Wirawan Adikusuma conceived and designed the study. Lalu Muhammad Irham, Wirawan Adikusuma, Syarifatul Mufidah, and Rahmat Dani Satria performed the computational analysis. Putri Permata Suka Admanegara, Wirawan Adikusuma, Darmawi Darmawi, Rockie Chong, and Lalu Muhammad Irham contributed to writing the manuscript. Lalu Muhammad Irham provided the funding. Lalu Muhammad Irham, Wirawan Adikusuma, Syarifatul Mufidah, Rahmat Dani Satria, Eko Mugiyanto, Darmawi Darmawi, Petrina Theda Philothra, Desti Rahmawati, Siska Widiyastuti, and Rista Yulianti contributed to the revision of the manuscript. Putri Permata Suka Admanegara, Brilliant Citra Wirashada, Sabiah Khairi, Indra Jaya, Baik Heni Rispawati, Rockie Chong, and Ilker Ates also revised the manuscript. All authors have read and approved the final version of the manuscript and have made significant contributions to the study.

References

- Admanegara PPS, Yulianti R, Rahmawati D, Widiastuti S, Adikusuma W, Wirashada BC, Amukti DP, Darmawi D, Rispawati BH, Iskandar B, *et al.* 2025. Leveraging Genome-wide Association Studies to Identify Pathogenic Variants for Breast Cancer Among Multiple Continents. *Anticancer Res.*, **45**: 5351-5367.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods.*, **7**: 248-249.
- Allemailem KS, Almatroudi A, Alrumaihi F, Makki Almansour N, Aldakheel FM, Rather RA, Afroze D and Rah B. 2021. Single nucleotide polymorphisms (SNPs) in prostate cancer: its implications in diagnostics and therapeutics. *Am J Transl Res.*, **13**: 3868-3889.
- Bıyıkoglu M, Tanrıverdi R, Bozlu M, Şenel S, Fidancı Ş B, Tamer L and Akbay E. 2024. Evaluation of homeobox protein B13 (HOXB13) gene G84E mutation in patients with prostate cancer. *World J Urol.*, **42**: 476.
- Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. 2024. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.*, **74**: 229-263.
- Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, *et al.* 2019. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**: D1005-D1012.
- Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, Netto GJ, Qin ZS, Kumar S, Manne U, *et al.* 2022. UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia.*, **25**: 18-27.
- Chatrath A, Ratan A and Dutta A. 2021. Germline Variants That Affect Tumor Progression. *Trends Genet.*, **37**: 433-443.
- Chu F, Chen L, Guan Q, Chen Z, Ji Q, Ma Y, Ji J, Sun M, Huang T, Song H, *et al.* 2025. Global burden of prostate cancer: age-period-cohort analysis from 1990 to 2021 and projections until 2040. *World J Surg Oncol.*, **23**: 98.
- Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, Field JR, Pulley JM, Ramirez AH, Bowton E, *et al.* 2013.

- Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol.*, **31**: 1102-1110.
- Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K and Liu X. 2015. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet.*, **24**: 2125-2137.
- Dupont WD, Breyer JP, Johnson SH, Plummer WD and Smith JR. 2021. Prostate cancer risk variants of the HOXB genetic locus. *Sci Rep.*, **11**: 11385.
- Egevad L, Granfors T, Karlberg L, Bergh A and Stattin P. 2002. Prognostic value of the Gleason score in prostate cancer. *BJU Int.*, **89**: 538-542.
- Erokhin M M, Kozelchuk NY, Ziganshin RH, Tatarskiy VV and Chetverina DA. 2025. HOXB13 interactome in prostate cancer cells: biochemical and functional interactions between the transcription factors HOXB13 and TBX3. *Vavilovskii Zhurnal Genet Selekt.*, **29**: 744-752.
- Frazer J, Notin P, Dias M, Gomez A, Min JK, Brock K, Gal Y and Marks DS. 2021. Disease variant prediction with deep generative models of evolutionary data. *Nature.*, **599**: 91-95.
- Gong H, Chen X, Jin Y, Lu J, Cai Y, Wei O, Zhao J, Zhang W, Wen X, Wang Y, *et al.* 2019. Expression of ARHGAP10 correlates with prognosis of prostate cancer. *Int J Clin Exp Pathol.*, **12**: 3839-3846.
- Halla Abdul-Hadi Chabuk ZIJ and Wathiq Mohammed.. 2025. Serotonin Gene Polymorphisms Association with Immune Response, Intestinal Inflammation, and Anxiety Behaviour in Rats Exposed to Sodium Arsenite. *Jordan J Biol Sci.*, **18**: 383-392.
- Heise M, Jarzowski P, Bąk A, Junkiert-Czarnecka A, Pilarska-Deltow M and Haus O. 2019. G84E germline mutation in HOXB13 gene is associated with increased prostate cancer risk in Polish men. *Pol J Pathol.*, **70**: 127-133.
- Huang KL, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, Paczkowska M, Reynolds S, Wyczalkowski MA, Oak N, *et al.* 2018. Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell.*, **173**: 355-370.e314.
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, *et al.* 2016. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet.*, **99**: 877-885.
- Kanayama M, Chen Y, Rabizadeh D, Vera L, Lu C, Nielsen SM, Russell EM, Esplin ED, Wang H, Isaacs WB, *et al.* 2024. Clinical and Functional Analyses of an African-ancestry Gain-of-function HOXB13 Variant Implicated in Aggressive Prostate Cancer. *Eur Urol Oncol.*, **7**: 751-759.
- Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM and Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.*, **46**: 310-315.
- Kote-Jarai Z, Mikropoulos C, Leongamornlert DA, Dadaev T, Tymrakiewicz M, Saunders EJ, Jones M, Jugurnauth-Little S, Govindasami K, Guy M, *et al.* 2015. Prevalence of the HOXB13 G84E germline mutation in British men and correlation with prostate cancer risk, tumour characteristics and clinical outcomes. *Ann Oncol.*, **26**: 756-761.
- Larkin SE, Holmes S, Cree IA, Walker T, Baskett V, Bickers B, Harris S, Garbis SD, Townsend PA and Aukim-Hastie C. 2012. Identification of markers of prostate cancer progression using candidate gene expression. *Br J Cancer.*, **106**: 157-165.
- Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skythe A and Hemminki K. 2000. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med.*, **343**: 78-85.
- Lin HY, Amankwah EK, Tseng TS, Qu X, Chen DT and Park JY. 2013. SNP-SNP interaction network in angiogenesis genes associated with prostate cancer aggressiveness. *PLoS One.*, **8**: e59688.
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, *et al.* 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science.*, **337**: 1190-1195.
- Ng PC and Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.*, **31**: 3812-3814.
- Ng PC and Henikoff S. 2001. Predicting deleterious amino acid substitutions. *Genome Res.*, **11**: 863-874.
- Ohuchida K, Mizumoto K, Ishikawa N, Fujii K, Konomi H, Nagai E, Yamaguchi K, Tsuneyoshi M and Tanaka M. 2005. The role of S100A6 in pancreatic cancer development and its clinical implication as a diagnostic marker and therapeutic target. *Clin Cancer Res.*, **11**: 7785-7793.
- Patel RA, Sayar E, Coleman I, Roudier MP, Hanratty B, Low JY, Jaiswal N, Ajkunic A, Dumpit R, Ercan C, *et al.* 2024. Characterization of HOXB13 expression patterns in localized and metastatic castration-resistant prostate cancer. *J Pathol.*, **262**: 105-120.
- Pritchard CC, Mateo J, Walsh MF, De Sarkar N, Abida W, Beltran H, Garofalo A, Gulati R, Carreira S, Eeles R, *et al.* 2016. Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. *N Engl J Med.*, **375**: 443-453.
- Reva B, Antipin Y and Sander C. 2011. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.*, **39**: e118.
- Shah R, Battisti NML, Brain E, Gnanon FHR, Kanesvaran R, Mohile S, Noronha V, Puts M, Soto-Perez-de-Celis E and Pilleron S. 2025. Updated cancer burden in oldest old: A population-based study using 2022 Globocan estimates. *Cancer Epidemiol.*, **95**: 102716.
- Sözen G.J.A.-H.a.E.. 2024. Determination of the Genetic Variations of Chickpea (*Cicer Arietinum* L.) Genotypes Preserved in the Jordanian Seed Genbank Using ISSR and SCot Molecular Markers. *Jordan J Biol Sci.*, **17**: 727-735.
- Srinivasan S, Kryza T, Bock N, Tse BWC, Sokolowski KA, Janathani P, Fernando A, Moya L, Stephens C, Dong Y, *et al.* 2024. A PSA SNP associates with cellular function and clinical outcome in men with prostate cancer. *Nat Commun.*, **15**: 9587.
- Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, *et al.* 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.*, **47**: D941-D947.
- Usha Adiga NH and Rajeev TP.. 2025. Evaluation of Gene Polymorphisms in Patients with Urinary Oxalate Stones: Cross-sectional Study. *Jordan J Biol Sci.*, **18**: 53-62.
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr. and Kinzler KW. 2013. Cancer genome landscapes. *Science.*, **339**: 1546-1558.
- Wang A, Shen J, Rodriguez AA, Saunders EJ, Chen F, Janivara R, Darst BF, Sheng X, Xu Y, Chou AJ, *et al.* 2023. Characterizing prostate cancer risk through multi-ancestry genome-wide discovery of 187 novel risk variants. *Nat Genet.*, **55**: 2065-2074.
- Witte JS, Mefford J, Plummer SJ, Liu J, Cheng I, Klein EA, Rybicki BA and Casey G. 2013. HOXB13 mutation and prostate cancer: studies of siblings and aggressive disease. *Cancer Epidemiol Biomarkers Prev.*, **22**: 675-680.

Zabalza CV, Adam M, Burdelski C, Wilczak W, Wittmer C, Kraft S, Krech T, Steurer S, Koop C, Hube-Magg C, *et al.* 2015. HOXB13 overexpression is an independent predictor of early PSA recurrence in prostate cancer treated by radical prostatectomy. *Oncotarget.*, **6**: 12822-12834.

Zhang J, Xiao L, Qin Z, Xu A, Zhao K, Liang C, Miao C, Zhu J, Chen W, Hua Y, *et al.* 2016. Association between germline homeobox B13 (HOXB13) G84E allele and prostate cancer susceptibility: a meta-analysis and trial sequential analysis. *Oncotarget.*, **7**: 67101-67110.