FEDERAL UNIVERSITY OF CEARÁ

CENTER OF AGRICULTURE SCIENCES

GRADUATE PROGRAM IN BIOTECHNOLOGY OF NATURAL RESOURCES

FRANCISCO JONATHAN DOS SANTOS ARAÚJO

CRIBRIFORM MORPHOLOGY IN PROSTATE CANCER: AN INTEGRATED APPROACH COMBINING DIGITAL PATHOLOGY AND TRANSCRIPTOMIC ANALYSIS

FORTALEZA

2025

FRANCISCO JONATHAN DOS SANTOS ARAÚJO

CRIBRIFORM MORPHOLOGY IN PROSTATE CANCER: AN INTEGRATED APPROACH COMBINING DIGITAL PATHOLOGY AND TRANSCRIPTOMIC ANALYSIS

Thesis submitted to the graduate program in biotechnology of natural resources of the center of agriculture sciences of the Federal University of Ceará, as a partial requirement for obtaining the title of Doctor in Biotechnology. Concentration Area: Biotechnology

Advisor: Prof. Dra. Vânia Maria Maciel Melo

Co-advisor: Dr. Fábio Fernandes Távora

FORTALEZA

2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A689c Araújo, Francisco Jonathan dos Santos.
Cribriform Morphology in Prostate Cancer: An Integrated Approach Combining Digital Pathology and Transcriptomic Analysis / Francisco Jonathan dos Santos Araújo. – 2025.
76 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Biotecnologia de Recursos Naturais, Fortaleza, 2025.
Orientação: Profa. Dra. Vânia Maria Maciel Melo.
Coorientação: Prof. Dr. Fábio Rocha Fernandes Távora.

1. Prostate Cancer. 2. Cribriform Morphology. 3. Prognostic Biomarkers. 4. Digital Pathology. 5. RNA-Seq analysis. I. Título.

CDD 660.6

FRANCISCO JONATHAN DOS SANTOS ARAÚJO

CRIBRIFORM MORPHOLOGY IN PROSTATE CANCER: AN INTEGRATED APPROACH COMBINING DIGITAL PATHOLOGY AND TRANSCRIPTOMIC ANALYSIS

Thesis submitted to the graduate program in biotechnology of natural resources of the center of agriculture sciences of the Federal University of Ceará, as a partial requirement for obtaining the title of Doctor in Biotechnology. Concentration Area: Biotechnology

Approved on: August 26, 2025

EXAMINATION BOARD

Prof. Dra. Vânia Maria Maciel Melo (Advisor) Universidade Federal do Ceará (UFC)	Dr. Fábio Fernandes Távora (Co-advisor) Universidade Federal do Ceará (UFC)
Profa. Dra. Raquel Carvalho Montenegro Universidade Federal do Ceará (UFC)	Dra. Luciana Schultz Amorim Instituto de Anatomia Patológica (iAP)
Profa. Dra. Adriane Feijó Evangelista Hospital de Amor de Barretos	Dra. Juliana Cordeiro de Sousa Sá Laboratório Argos Patologia

I dedicate this thesis to my mother, Francisca Sandra, for her unconditional love, prayers, and for being my strength in the most difficult moments. And to my advisor, Dr. Vânia Maria Maciel Melo, my scientific mother, for her inspiration and generosity.

ACKNOWLEDGEMENTS

To Professor Vânia Maria Maciel Melo, my advisor, for her constant presence, unwavering support, and inspiration. Thank you for being my scientific mother and a true example of professionalism, ethics, and dedication to research.

To Professor Dr. Fábio Fernandes Távora, my mentor, for all the support and for introducing me—passionately and with excellence—to the fascinating world of pathology and cancer research.

To Dra. Juliana Cordeiro, a dear friend and inspiring leader, for all the support, friendship, and encouragement during my immersion into the world of molecular biology applied to health and oncology. Thank you also for the laughter and for so many valuable lessons.

To my bioinformatics mentor, Adriane Feijó, for her support, thoughtful guidance, and constant encouragement, which were essential throughout the course of this thesis.

To my great friend Guilherme Velozo, for the work partnership, daily dedication and affection, and for the essential support in applying artificial intelligence to our projects. Your friendship and warm hugs made all the difference.

To my friends Emerson and Jamilly, from the molecular biology team at Argos Laboratory, for their technical support, friendship, dedication, and companionship in every stage of this work.

To the entire Argos Laboratory team, for their warmth, support, shared laughter, and all the knowledge exchanged along the way.

To the IPEC team in Guarapuava-PR, for welcoming me so generously and for all the learning and meaningful exchanges during my time as a Genoma SUS fellow.

To my friends and soul siblings in Guarapuava — Robert Costa, Caroline Viana, Regina Maurina, Vitor Gregório, Maria Holzer, and Ana Cristina — for all the joy, companionship, and friendship that made this chapter lighter and unforgettable.

To my love, Ian Lima, for standing by my side with affection, support, and dedication, helping make this dream come true and building our life journey together.

To my mother, Francisca Sandra, for her unconditional love, faith, and all the prayers that sustained me through the hardest moments.

To God, for His divine guidance, strength, and grace that sustained me through every challenge; and to my beloved family, whose unconditional love, constant support, and prayers were a source of light and encouragement along this journey.

“Eu sou de uma terra que o povo padece, mas não esmorece e procura vencer.”

(Patativa do Assaré)

ABSTRACT

Cribriform morphology has been increasingly recognized as a histological marker of aggressiveness in prostate cancer, being associated with a higher risk of recurrence, metastasis, and cancer-specific mortality. However, the molecular mechanisms underlying this architectural pattern remain poorly understood. This thesis proposes a multidimensional approach to investigate the morphological and molecular determinants associated with tumor aggressiveness in prostate cancer, with a particular focus on cribriform architecture. In the first chapter, 72 radical prostatectomy samples were analyzed using quantitative digital pathology. Through whole-slide imaging and the QuPath software, immunohistochemical markers PTEN, Ki-67, ATM, and CD8 and key morphological features—such as Gleason pattern 4 percentage, tumor extent, and cribriform morphology—were quantified. The analysis identified a group of markers associated with increased tumor aggressiveness, with cribriform morphology showing significant correlation with higher pathological stage and absence of organ confinement. A logistic regression model combining Gleason pattern 4 percentage and ATM expression accurately predicted the presence of cribriform morphology (AUC = 0.79). In the second chapter, an in silico RNA-Seq analysis was conducted using publicly available data (GEO-NCBI) to characterize the gene expression profile associated with cribriform morphology. A total of 55 differentially expressed genes were identified between cribriform and non-cribriform tumors. Genes such as *CPNE4, PLA2G7, PTPRJ, GCNT1, AMACR, AR,* and *ERG* were linked to immune evasion, lipid metabolism, and hormonal signaling pathways. Functional enrichment analysis revealed transcriptional reprogramming consistent with aggressive tumor behavior. Although methodologically independent, the two chapters present complementary findings that reinforce the association between cribriform morphology and a biologically aggressive phenotype in prostate cancer, contributing to the identification of prognostic biomarkers and potential therapeutic targets

Keywords: prostate cancer; cribriform morphology; digital pathology; gene expression; prognostic biomarkers.

RESUMO

A morfologia cribriforme tem sido cada vez mais reconhecida como um marcador histológico de agressividade no câncer de próstata, estando associada a maior risco de recorrência, metástase e mortalidade específica pela doença. No entanto, os mecanismos moleculares que sustentam esse padrão arquitetural ainda são pouco compreendidos. Esta tese propõe uma abordagem multidimensional para investigar os determinantes morfológicos e moleculares associados à agressividade tumoral no câncer de próstata, com foco especial na arquitetura cribriforme. No primeiro capítulo, foram analisadas 72 amostras de prostatectomia radical utilizando patologia digital quantitativa. Por meio da digitalização de lâminas inteiras e do uso do software QuPath, foram quantificados marcadores imunohistoquímicos PTEN, Ki-67, ATM e CD8, além de características morfológicas-chave, como o percentual de padrão 4 de Gleason, extensão tumoral e presença de morfologia cribriforme. A análise identificou um conjunto de marcadores associados ao aumento da agressividade tumoral, sendo a morfologia cribriforme significativamente correlacionada com estadiamento patológico mais avançado e ausência de confinamento ao órgão. Um modelo de regressão logística, combinando o percentual de padrão 4 com a expressão de ATM, previu com precisão a presença da morfologia cribriforme (AUC = 0,79). No segundo capítulo, foi realizada uma análise in silico de dados de RNA-Seq provenientes de banco público (GEO-NCBI), com o objetivo de caracterizar o perfil de expressão gênica associado à morfologia cribriforme. Foram identificados 55 genes diferencialmente expressos entre tumores cribriformes e não cribriformes. Genes como *CPNE4, PLA2G7, PTPRJ, GCNT1, AMACR, AR* e *ERG* estiveram associados a vias de evasão imune, metabolismo lipídico e sinalização hormonal. A análise funcional indicou uma reprogramação transcricional compatível com comportamento tumoral agressivo. Embora metodologicamente independentes, os dois capítulos apresentam achados complementares que reforçam a associação entre a morfologia cribriforme e um fenótipo biologicamente agressivo no câncer de próstata, contribuindo para a identificação de biomarcadores prognósticos e potenciais alvos terapêuticos.

Palavras-chave: câncer de próstata; morfologia cribriforme; patologia digital; expressão gênica; biomarcadores prognósticos.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AMACR	*Alpha-Methylacyl-CoA Racemase*
ANKRD30A	*Ankyrin Repeat Domain 30A*
AR	*Androgen Receptor*
CPNE4	*Copine 4*
DOLPP1	*Dolichyldiphosphatase 1*
ERG	*ETS-Related Gene*
FAM124A	*Family With Sequence Similarity 124 Member A*
GCNT1	*Glucosaminyl (N-Acetyl) Transferase 1*
LRRC61	*Leucine Rich Repeat Containing 61*
NAALADL2	*N-Acetylated Alpha-Linked Acidic Dipeptidase Like 2*
NGEF	*Neuronal Guanine Nucleotide Exchange Factor*
PLA2G7	*Phospholipase A2 group VII*
POSTN	*Periostin*
PTPRJ	*Protein Tyrosine Phosphatase Receptor Type J*
SLC13A4	*Solute Carrier Family 13 Member 4*
TMPRSS5	*Transmembrane Protease, Serine 5*
TRIM9	*Tripartite Motif Containing 9*
ATM	Ataxia Telangiectasia Mutated
AUC	Area Under the Curve
BP	Biological Process
CD8	$CD8^+$ T lymphocytes
CI	Confidence Interval
CP	Cribriform Pattern
DAB	Diaminobenzidine Chromogen
DEA	Differential Expression Analysis
DEGs	Differentially Expressed Genes
DNA	Deoxyribonucleic Acid
FFPE	Formalin-Fixed Paraffin-Embedded
GEO	Gene Expression Omnibus
GO	Gene Ontology
GP4	Gleason Pattern 4

H&E	Hematoxylin and Eosin
H-score	Histological Score
HR	Hazard Ratio
ICC	Invasive Cribriform Carcinoma
IDC	Intraductal Carcinoma
ISUP	International Society of Urological Pathology
KI-67	KI-67 Antigen
LogFC	Log Fold Change
MF	Molecular Function
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing
OR	Odds Ratio
PC	Prostate Cancer
PSA	Prostate-Specific Antigen
PTEN	Phosphatase and Tensin Homolog
RNA-seq	RNA Sequencing
ROC	Receiver Operating Characteristic
RP	Radical Prostatectomy
TIL	Tumor-Infiltrating Lymphocyte
TMA	Tissue Microarray
WSI	Whole Slide Imaging

LIST OF SYMBOLS

mM	Millimolar (concentration)
mg/mL	Milligrams per milliliter (concentration)
ng	Nanograms (mass)
pH	Hydrogen potential (acidity/alkalinity)
°C	Degrees Celsius (temperature)
mmHg	Millimeters of mercury (pressure)
mm	Millimeters (length)
μm	Micrometers (length)
μm^2	Square micrometers (area)
cells/mm^2	Cell density (cells per square millimeter)
n	Sample size or number of observations
%	Percentage
p	p-value (statistical significance)

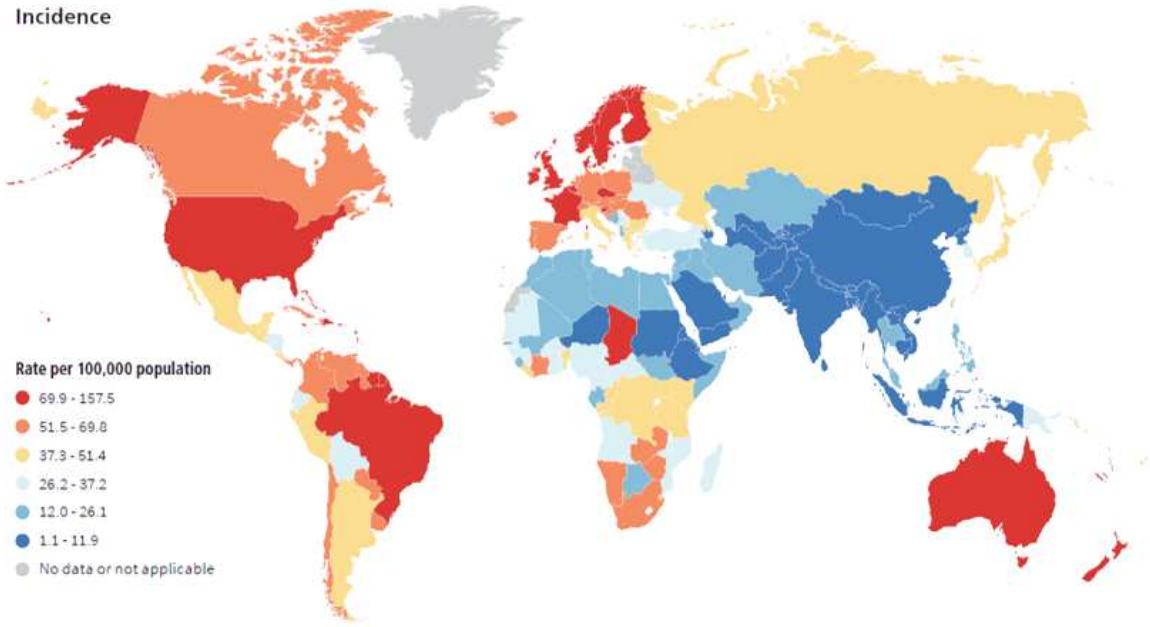
CONTENTS

1 INTRODUCTION

1.1 Prostate cancer

Prostate cancer (PC) is the second most common type of cancer among men worldwide and one of the leading causes of cancer-related mortality in both developed and developing countries (Rawla *et al.*, 2022). In 2020, approximately 1.4 million new global cases were reported, with an incidence of 78 per 100,000 inhabitants, making it the most frequently diagnosed cancer among men across all regions of Brazil (Sung *et al.*, 2021; INCA, 2022). This global distribution is illustrated in Figure 1, which shows prostate cancer as the most frequently diagnosed cancer in men across 118 countries. Brazil is notably among the top 10 countries in incidence, underscoring its significant public health impact and aligning with national epidemiological data that place the disease as a major oncological concern.

Figure 1 – Prostate cancer distribution worldwide



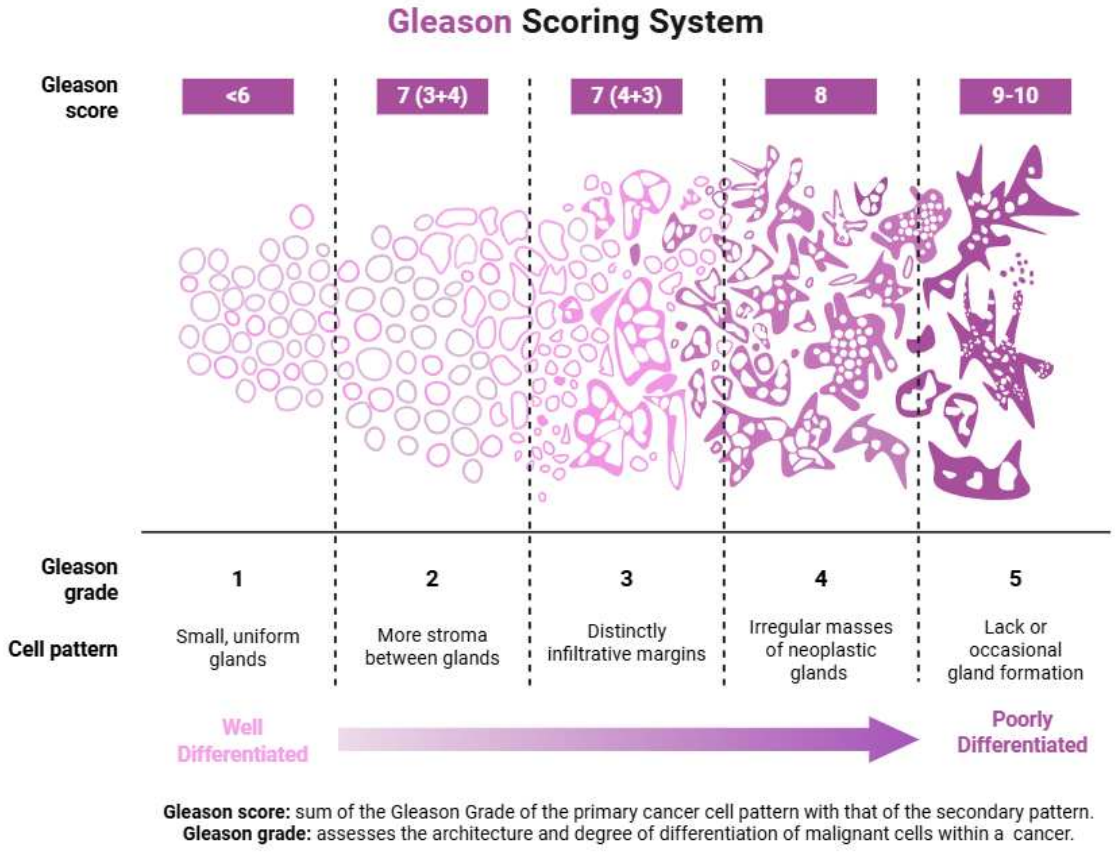
Source: American cancer society, 2025. Available at: <https://www.cancer.org/research/acs-research-news/prostate-cancer-is-number-1-for-118-countries-worldwide.html>. Accessed on: Aug. 2, 2025.

It is a notoriously heterogeneous disease, with multiple molecular profiles that remain poorly understood (Haffner *et al.*, 2021; Barbieri *et al.*, 2021). The main risk factors include age, family history, and ethnicity, with age being the only universally recognized risk factor (Cui *et al.*, 2024). Diagnosis involves assessing PSA levels, digital rectal examination, and imaging tests such as magnetic resonance imaging and ultrasound, although definitive confirmation relies

on histopathological analysis (Haffner *et al.*, 2021). Clinical risk stratification is based on the combination of PSA, Gleason score, and tumor staging, but these variables still have limitations in predicting tumor biological behavior (Richards *et al.*, 2022).

The Gleason scoring system ranges from 2 to 10, with higher scores reflecting poorly differentiated tumors and a strong correlation with aggressive clinical behavior (Mun *et al.*, 2021). The Gleason score is determined by adding the two most prevalent histological patterns: the primary pattern (predominant architecture) and the secondary pattern (second most common architecture), and the sum represents the Gleason score (Van Leenders *et al.*, 2020).

Figure 2 – Illustration of Gleason grading patterns based on the degree of prostate tumor cell differentiation



Source: Adapted from Yang *et al.*, 2022. Available at: <https://doi.org/10.4236/ss.2022.138047>. Accessed on: Aug. 2, 2025

In 2005, the International Society of Urological Pathology (ISUP) introduced significant modifications to the Gleason grading system, refining the definitions of individual patterns. These changes led to a substantial reclassification of tumors, with a marked reduction in Gleason score 6 cases and a corresponding increase in Gleason score 7 diagnoses. For instance, a biopsy previously scored as 3 + 3 = 6 with a minor (<5%) component of pattern 4 would now be graded

as 3 + 4 = 7 (Short, Warren e Varma, 2019).

Following these revisions, Helpap *et al.* (2016) reported a decline in Gleason 6 cases from 48% to 22% and an increase in Gleason 7 cases (3 + 4 and 4 + 3) from 26% to 68%. While patients with Gleason 6 tumors are generally considered low risk, Gleason 7 tumors frequently warrant more aggressive interventions such as radical prostatectomy or radiotherapy (Hollemans *et al.*, 2019). Despite these refinements, risk stratification within Gleason 7 remains challenging, emphasizing the need for additional prognostic markers to improve clinical decision-making.

In 2014, the ISUP and the World Health Organization (WHO) endorsed the prognostic grade group system, derived from Gleason scores and validated across multiple academic centers based on biochemical recurrence rates (Epstein *et al.*, 2016). The classification of the five prognostic grade groups is detailed in Table 1.

Table 1 – Classification of Grade Groups according to Gleason score.

Grade Group	**Gleason Score**	**Risk Category**
1	≤ 6	Low-grade PCa
2	3 + 4 = 7	Intermediate grade
3	4 + 3 = 7	Worse prognosis than Group 2
4	8	High-grade PCa
5	9–10	High-grade PCa

Source: Prepared by the author (2025).
Note: Adapted from ISUP classification.

Within Gleason Grade 4, four distinct architectural patterns are recognized—poorly formed, fused, glomeruloid, and cribriform. Of these, the cribriform pattern stands out as the most morphologically and clinically heterogeneous, with strong associations with adverse outcomes (Kweldam *et al.*, 2018).

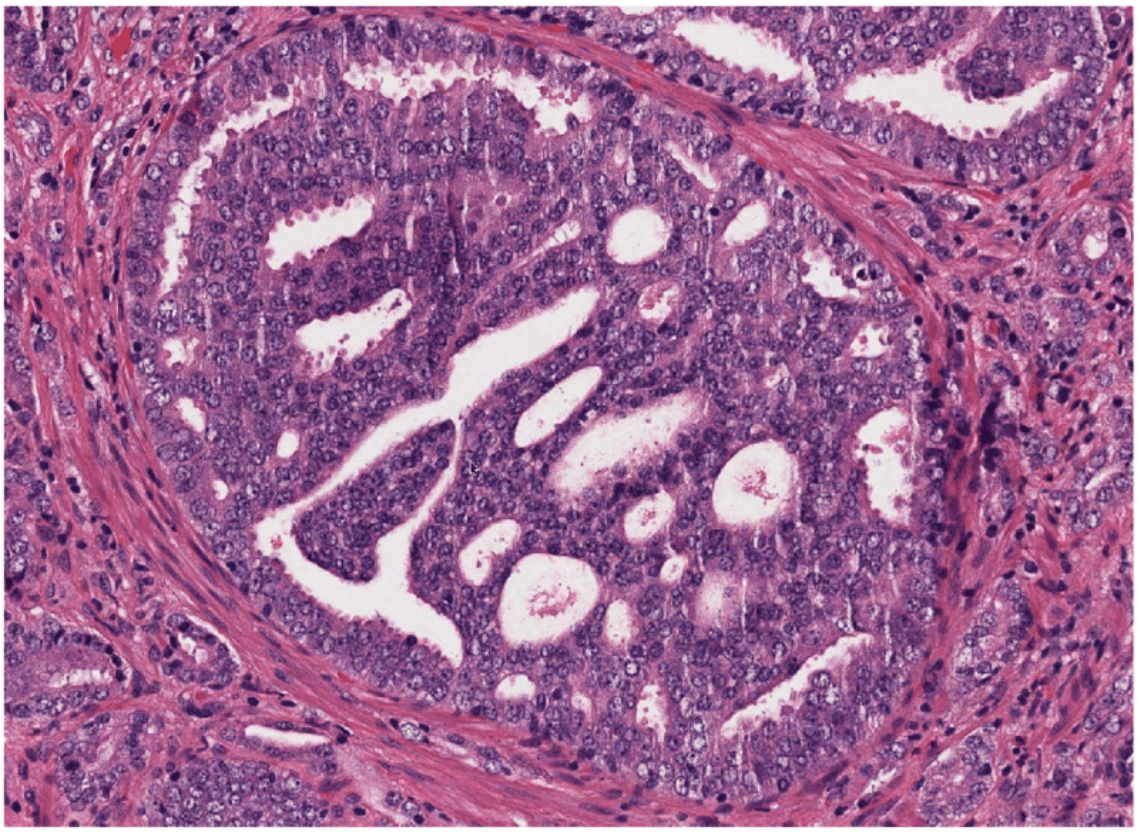
1.2 Cribriform pattern

Among the architectural subtypes of Gleason Grade 4, the cribriform pattern (CP) has emerged as a morphologic hallmark of aggressive prostate cancer. Its clinical relevance has been progressively recognized over the past two decades, following the 2005 International Society of Urological Pathology (ISUP) consensus, which established that any cribriform architecture should be graded as pattern 4, regardless of its extent (Epstein *et al.*, 2016; Kweldam *et al.*, 2018).

Histologically, CP is defined by epithelial proliferations forming large, irregularly

contoured glandular spaces with punched-out lumina and lacking intervening stroma (Kweldam *et al.*, 2018) (Figure 3). CP may occur in invasive or intraductal forms, both of which are associated with adverse prognosis, although intraductal carcinoma (IDC-P) typically exhibits even stronger associations with advanced stage and poor outcomes (Hollemans *et al.*, 2019).

Figure 3 – Example of cribriform pattern in prostate cancer samples



Source: Lotan *et al.*, 2017. Available at: <https://doi.org/10.48550/arXiv.1705.02678>. Accessed on: Aug. 2, 2025.

A growing body of evidence demonstrates that CP is independently associated with adverse pathological and clinical outcomes. Multiple retrospective and prospective studies have shown that the presence of CP in radical prostatectomy or biopsy specimens correlates with higher pathological stage, extraprostatic extension, seminal vesicle invasion, lymph node metastasis, and biochemical recurrence (Hollemans *et al.*, 2019; Zhang *et al.*, 2022). Notably, Zhang *et al.* (2022) reported that CP was associated with a nearly twofold increased risk of cancer-specific mortality after adjusting for PSA, tumor stage, and Gleason score.

Beyond its clinical correlations, CP is increasingly recognized as a distinct biological entity within prostate cancer. Molecular analyses have demonstrated that CP tumors exhibit higher frequencies of PTEN loss, *TMPRSS2-ERG* fusion, and genomic instability compared

to non-cribriform tumors of the same grade group (Bottcher *et al.*, 2018; Van Leenders *et al.*, 2020). Transcriptomic studies further suggest a unique gene expression profile in CP tumors, characterized by dysregulation of pathways involved in cell adhesion, extracellular matrix remodeling, lipid metabolism, and immune evasion (Song *et al.*, 2022). These findings support the hypothesis that CP represents not merely a morphological variant, but a biologically aggressive phenotype driven by distinct molecular mechanisms.

Despite its clinical and biological significance, CP recognition in routine diagnostic practice remains largely subjective. Interobserver variability in identifying CP—especially in small foci or in borderline cases with fused glands—remains a challenge, even among experienced genitourinary pathologists (Van der slot *et al.*, 2020; Kweldam *et al.*, 2018). This subjectivity limits reproducibility and may contribute to inconsistencies in clinical reporting and research outcomes.

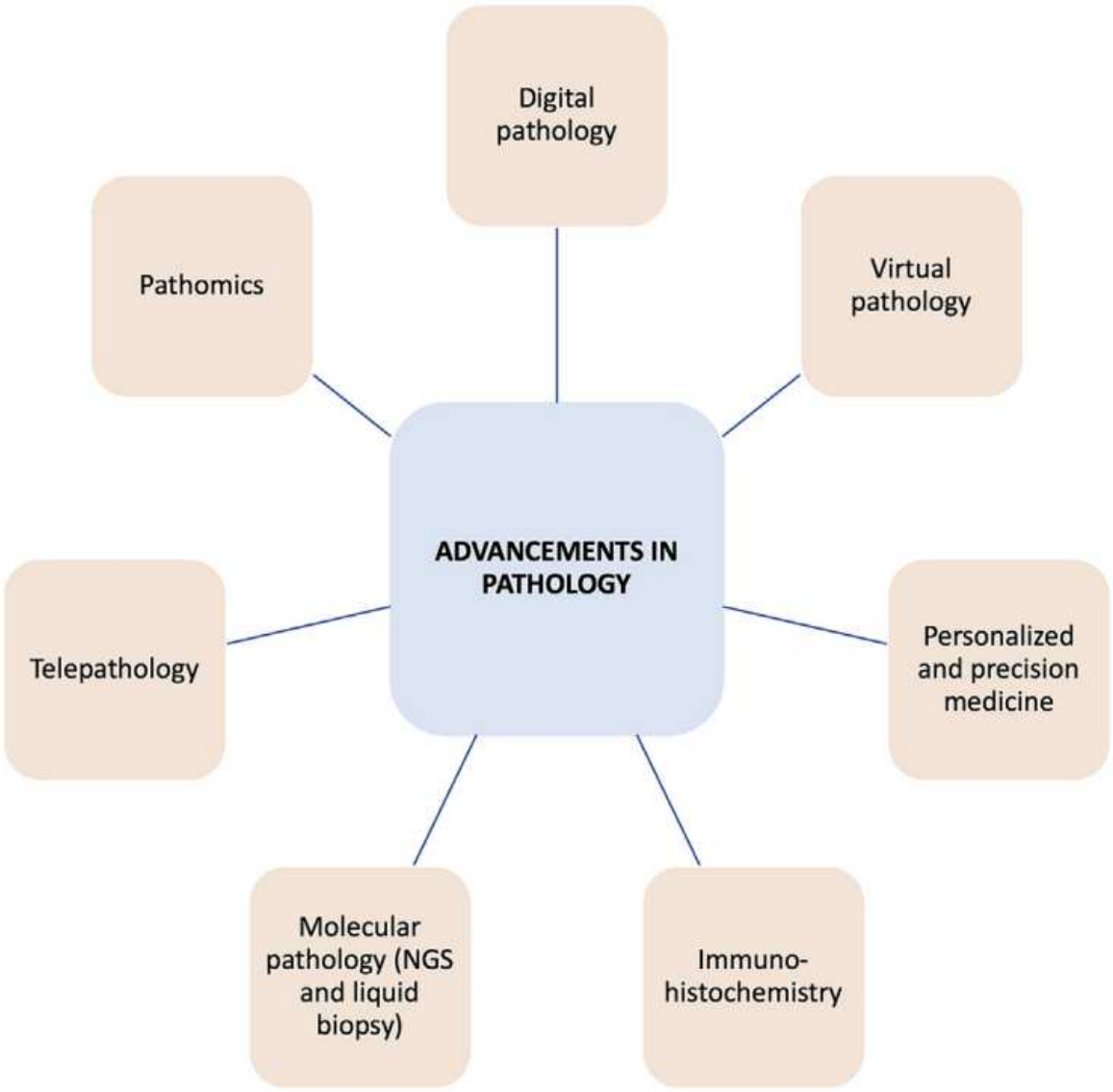
Given these limitations, there is a growing interest in developing objective, reproducible tools to evaluate CP. Advances in digital pathology and computational image analysis offer promising opportunities to standardize CP detection and quantification. Moreover, integrating morphological evaluation with molecular characterization may refine risk stratification and facilitate the identification of novel prognostic biomarkers and therapeutic targets.

1.3 Digital Pathology

Digital pathology (DP) has emerged as a transformative tool in modern histopathology, enabling greater standardization, scalability, and precision in the evaluation of histological specimens (Rabilloud *et al.*, 2023; Sirinukunwattana *et al.*, 2021). The digitization of glass slides into high-resolution whole-slide images (WSIs) facilitates remote consultation, large-scale archiving, and advanced computational analysis (Barker *et al.*, 2024). As illustrated in Figure 4, DP is part of a broader digital transformation in pathology that encompasses virtual pathology, telepathology, molecular approaches (including next-generation sequencing and liquid biopsy), immunohistochemistry, and personalized medicine, all of which contribute to more integrated and data-driven diagnostic workflows.

Whole-slide imaging (WSI) refers to the scanning of entire histological slides at high magnification—typically 20× or 40×—producing high-resolution digital files that may reach several gigabytes per slide (Clarke *et al.*, 2023). WSIs faithfully capture all visual details available through traditional microscopy, enabling pathologists and computational tools to

Figure 4 – Overview of advancements in pathology enabled by digital transformation, including digital and virtual pathology, telepathology, molecular approaches, immunohistochemistry, and personalized medicine



Source: Barker *et al.*, 2024. Available at: <https://doi.org/10.1016/j.jpi.2024.100408>. Accessed on: Aug. 2, 2025

navigate specimens at multiple zoom levels. As illustrated in Figure 5, the WSI workflow involves a sequence of steps from slide selection and image acquisition to digital analysis, data storage, and telepathology. This process not only supports local visualization but also forms the foundation for remote pathology practice, facilitating second opinions, intraoperative consultations, and broader collaboration through standardized digital formats, including DICOM-WSI for improved interoperability.

Despite its benefits, the implementation of DP faces technical and infrastructure challenges. The acquisition of high-quality scanners, the storage of large datasets, and the integration with laboratory information systems (LIS) require substantial investment (Lee *et al.*,

Figure 5 – Workflow of whole-slide imaging (WSI) from slide selection to telepathology.

Source: Barker *et al.*, 2024. Available at: <https://doi.org/10.1016/j.jpi.2024.100408>. Accessed on: Aug. 2, 2025

2024). WSIs demand robust computational resources, with significant memory requirements, high-performance network bandwidth for real-time navigation, and long-term digital storage solutions (Basak *et al.*, 2023). Additionally, variability among scanners, staining protocols, and image compression methods can affect data quality and standardization (Clarke *et al.*, 2023).

Artificial intelligence (AI) has become an essential complement to DP, expanding its potential from visualization to advanced diagnostic support. AI-based models—especially those using deep learning—can segment tissue compartments, classify tumor patterns, quantify nuclear features, and identify histologic patterns such as the cribriform pattern (CP) wi th high accuracy (Asif *et al.*, 2023). In prostate cancer, digital pathology has shown particular value in the objective assessment of aggressive histological patterns such as the cribriform pattern (CP). Computational image analysis enables precise identification and quantification of cribriform pattern areas, evaluation of nuclear morphology, glandular architecture, and biomarker expression. When combined with artificial intelligence (AI), these approaches achieve performance comparable to experienced pathologists in grading, while substantially reducing interobserver variability (Singhal *et al.*, 2022; Bulten *et al.*, 2020). AI-driven quantification of CP, integrated with other histological and molecular parameters, has been applied to large datasets to validate CP as an independent prognostic factor and to develop predictive models for progression, recurrence, and mortality (Zhang *et al.*, 2023).

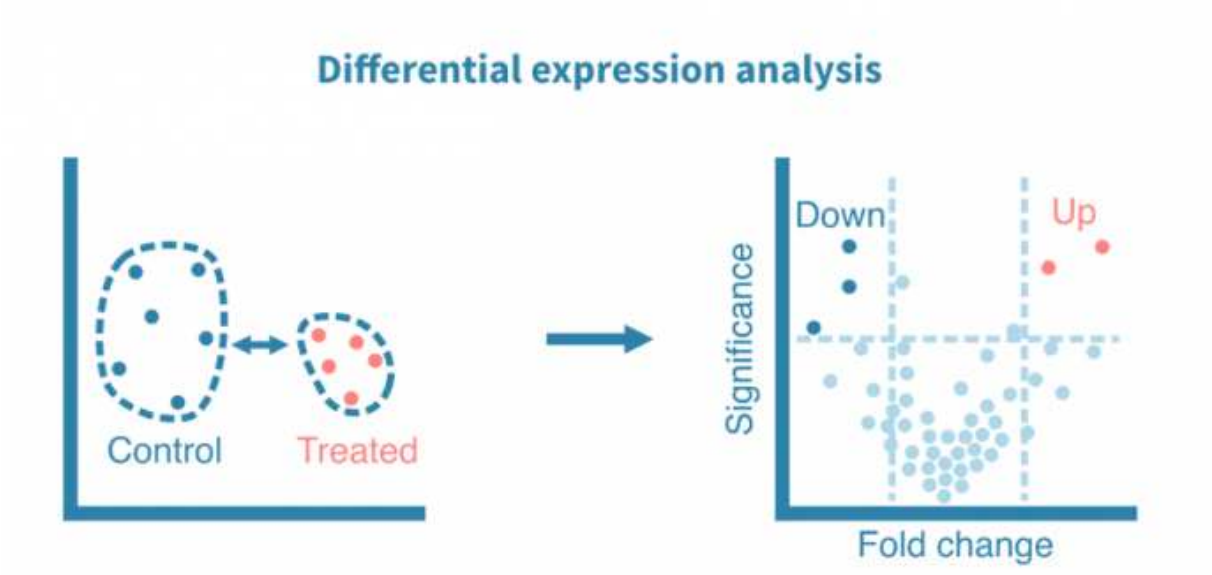
Looking forward, DP, integrated with AI and multi-omic data, represents a bridge between traditional histopathology and precision oncology. By combining morphological features extracted from WSIs with transcriptomic, proteomic, and genomic profiles, it will be possible to identify biologically relevant patterns and support more accurate prognostic modeling in prostate cancer.

1.4 Transcriptomics

In addition to morphological analysis, transcriptomic profiling has become an essential strategy for elucidating the molecular mechanisms underlying tumor heterogeneity (Feng *et al.*, 2024). Transcriptomics refers to the comprehensive evaluation of RNA transcripts expressed in a given sample, providing a dynamic view of cellular activity that reflects both genetic alterations and microenvironmental influences (Salachan *et al.*, 2023).

Among the available techniques, RNA sequencing (RNA-seq) has emerged as the gold standard for transcriptome analysis (Corchete *et al.*, 2020). When combined with robust bioinformatic pipelines, RNA-seq enables the identification of differentially expressed genes (DEGs) between biologically or clinically relevant groups—such as tumors with and without specific morphological patterns (De-Chao *et al.*, 2024) (Figure 6). These analyses reveal gene expression signatures that may underlie phenotypic differences, offering insights into pathways, regulatory networks, and biological processes associated with tumor aggressiveness (Wong *et al.*, 2022).

Figure 6 – The concept of differential gene expression, in which upregulated and downregulated transcripts are systematically evaluated to identify biologically relevant changes between experimental groups.



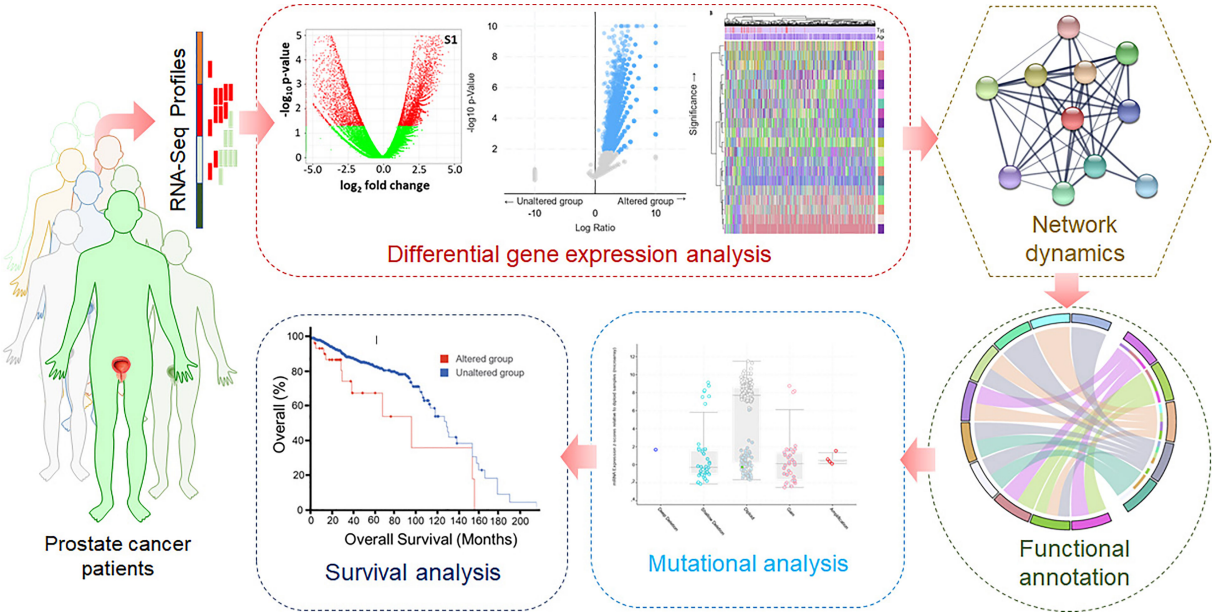
Source: Genvia technologies, s.d. Available at: <https://geneviatechnologies.com/bioinformatics-analyses/rna-seq-data-analysis/>. Accessed on: Aug. 2, 2025

In prostate cancer, the cribriform pattern (CP) provides a particularly relevant clinical and histological model for transcriptomic investigation due to its strong association with adverse

pathological features and poor clinical outcomes. Comparative transcriptomic studies have shown that CP tumors often exhibit deregulation of cell adhesion molecules, extracellular matrix remodeling, lipid metabolism, and immune-related pathways (Ma *et al.*, 2023). Such molecular features are consistent with the distinct morphology and aggressive behavior observed histologically.

Bioinformatic analysis of differential expression genes (DEGs) typically involves multiple steps: normalization of expression counts, statistical modeling of differential expression (commonly using packages such as DESeq2, edgeR, or limma-voom), and functional enrichment analysis to identify overrepresented Gene Ontology (GO) terms or Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Costa-Silva *et al.*, 2021; Paton *et al.*, 2024). These downstream analyses allow for the integration of transcriptomic signatures with biological interpretation, enabling the discovery of potential biomarkers and therapeutic targets, as illustrated in Figure 7 (Wang *et al.*, 2022).

Figure 7 – Workflow of RNA-seq data analysis in prostate cancer, starting from transcriptomic profiling and proceeding through differential gene expression analysis. This integrative approach enables the identification of biologically relevant pathways and potential prognostic markers



Source: Wang *et al.*, 2022. Available at: <https://doi.org/10.3389/fonc.2022.881246>. Accessed on: Aug. 2, 2025

In the present study, differential expression analysis was performed using the edgeR package, a widely validated tool for the statistical analysis of RNA-seq count data. edgeR employs negative binomial models to account for variability in count data, providing accurate dispersion estimates even for datasets with limited replicates (Robinson *et al.*, 2010). This

approach is particularly advantageous in studies with moderate sample sizes, as it stabilizes variance estimates and controls the false discovery rate (FDR) through robust multiple testing correction. Furthermore, edgeR offers flexibility for complex experimental designs, allowing the incorporation of covariates and contrasts that accurately reflect the biological comparisons of interest.

Despite its utility, differential expression analysis (DEA) presents inherent limitations that must be considered in both design and interpretation. One major challenge is sample size imbalance between comparison groups, which can reduce statistical power and increase variance estimates, potentially leading to false negatives or inflated effect sizes (Schurch *et al.*, 2016). In the context of prostate cancer studies, where specific morphological subtypes such as cribriform pattern may be underrepresented, group size imbalance can compromise the robustness of DEG detection. Additional factors—such as batch effects, RNA quality, and biological heterogeneity—may further confound results. These limitations underscore the importance of validation in independent cohorts and through orthogonal methods (e.g., quantitative PCR or protein-level confirmation) to ensure the reproducibility and biological relevance of identified gene signatures.

The application of these approaches to public RNA-seq datasets, such as The Cancer Genome Atlas (TCGA) and other prostate cancer transcriptomic cohorts, provides an invaluable resource for investigating the molecular basis of CP. Integrating these data with digital pathology–derived morphological parameters represents a promising strategy to build comprehensive morpho-molecular models, capable of refining risk stratification and guiding precision oncology in prostate cancer.

2 OBJECTIVE

In light of this evidence, this thesis aims to investigate the morphological, molecular, and transcriptomic determinants associated with tumor aggressiveness in PC, with a particular focus on cribriform morphology. To achieve this, a multidimensional approach was adopted, integrating quantitative digital pathology and gene expression analysis.

3 METHODOLOGICAL APPROACH

The work was structured into two individual articles, with distinct methodological designs but conceptually complementary, presented in the form of scientific studies. Together, these chapters explore, from different perspectives, the factors involved in the biology of cribriform tumors and their prognostic implications.

The first article presents the analysis of 72 Radical Prostatectomy (RP) samples processed into TMA, with whole-slide scanning and automated quantification using the QuPath software. It evaluated the expression of markers such as Phosphatase and tensin homolog (PTEN), Ki-67 antigen (Ki-67), Ataxia telangiectasia mutated (ATM) e $CD8^+$ T lymphocytes (CD8), as well as key histological features, including GP4 percentage, tumor extent, and the presence of CP. The data were correlated with clinical and pathological variables, and statistical models were applied to identify patterns associated with tumor aggressiveness. The results revealed the formation of a group of markers correlated with an aggressive phenotype, in which CP played a central role. Additionally, a logistic regression model based on GP4 percentage and ATM expression showed good predictive performance for the presence of CP (AUC = 0.79), demonstrating the potential of digital pathology for identifying high-risk tumors.

The second article complements this analysis by exploring the molecular mechanisms underlying CP through an *in silico* approach. Transcriptomic data from publicly available samples were obtained from the Gene Expression Omnibus (GEO)-National Center for Biotechnology Information (NCBI) database, comparing cases with and without CP. After normalization and sample bias control using Euclidean distance-based strategies, 55 DEGs were identified between the groups. Genes such as *Copine 4* (*CPNE4*), *Phospholipase A2 group VII* (*PLA2G7*), and *Protein Tyrosine Phosphatase Receptor Type J* (*PTPRJ*) were found in both analyses and were associated with immune evasion, lipid metabolism, and hormonal response pathways, suggesting a pro-tumorigenic transcriptional profile in cribriform tumors. Furthermore, underexpressed genes such as *Tripartite Motif Containing 9* (*TRIM9*), *Homeobox D3* (*HOXD3*) and *Integrin Subunit Beta 3 Binding Protein* (*ITGB3BP*) pointed to a possible loss of regulatory functions in the tumor microenvironment. Functional analysis revealed transcriptional reprogramming consistent with increased biological aggressiveness, reinforcing the role of CP as an integrative marker of phenotypic and molecular alterations.

The integration of the two articles proposed in this thesis provides a comprehensive perspective on aggressiveness in PC by combining the objectivity of digital pathology with the

depth of molecular analysis. By bringing together morphological, immunohistochemical, and transcriptomic data, this work contributes to advancing the understanding of the factors that drive aggressive tumor behavior and proposes tools that may aid in prognostic stratification and clinical decision-making. The thesis is structured in a publication-based format, composed of two interdependent scientific chapters, followed by a final integrative discussion and concluding remarks.

4 QUANTITATIVE DIGITAL PATHOLOGY REVEALS MORPHOLOGICAL AND MOLECULAR CORRELATES OF TUMOR AGGRESSIVENESS IN PROSTATE CANCER

Francisco Araujo[1], Guilherme Velozo[2], Juliana Cordeiro[2], Aline Ramos[2], Samuel Ferreira[2], Laura Cardoso[2], Vania Melo[1], Fábio Távora[2]

[1]Biology Department, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil

[2]Laboratório Argos Patologia, Av. Santos Dummond, Fortaleza, Ceará, Brazil

Submission date: July 8, 2025

Journal submitted: International Urology and Nephrology

Status of submission: Under review

Abstract

Digital pathology offers new opportunities for the objective assessment of prognostic markers in prostate cancer through integrated morphological and molecular analysis. In this study, 72 radical prostatectomy samples were analyzed using tissue microarrays and whole-slide imaging. Expert pathologists annotated tumor regions and analyzed them with QuPath software to quantify the expression of PTEN, Ki-67, ATM and CD8, as well as key morphological characteristics. Digital quantification of PTEN expression revealed high variability between samples, reflecting tumor heterogeneity, but showed limited ability to distinguish PTEN loss (AUC = 0.61). The correlation analysis revealed a cluster of aggressiveness markers, with cribriform morphology integrated into this profile. PTEN showed an unexpected correlation with Ki-67, while CD8+ density supported an immunologically "cold" phenotype. Cribriform morphology was also significantly associated with a higher pathological stage and absence of organ confinement, strengthening its prognostic value. However, the quantitative cribriform area did not correlate with these features. Finally, a logistic regression model that combined the percentage of pattern 4 and ATM expression predicted cribriform morphology with good discriminative performance (AUC = 0.79). These findings support the utility of digital pathology in identifying relevant prognostic patterns and developing predictive tools for high-risk tumor characteristics in prostate cancer.

Keywords: Digital Pathology, Prostate Cancer, Cribriform Morphology, Prognostic Biomarkers.

4.1 Introduction

Digital pathology has significantly transformed the analysis of tumor tissues and the interpretation of immunohistochemical markers, promoting a shift from conventional pathology to a more objective and reproducible approach (Hijazi *et al.*, 2024). This technology is based on the digitization of conventional histological slides, which generate high-resolution images known as Whole Slide Imaging (WSI), enabling the acquisition, storage, and computational analysis of pathological data (Lopes *et al.*, 2024). In the context of PC, specialized algorithms, including deep learning-based tools, have been incorporated into clinical practice, aiding in automated biomarker quantification and detection of suspicious cases, with promising implications for diagnosis, prognosis, and development of theranostic strategies (He *et al.*, 2023).

Prostate cancer is currently recognized as a notoriously heterogeneous neoplasm, both morphologically and molecularly, presenting a challenge to clinical management and prognostic stratification (Paralkar *et al.*, 2024). Although multiple risk factors have been identified, including age, family history, and ethnicity, advanced age remains the only well-established and widely accepted risk factor in the literature (Vaidyanathan *et al.*, 2016; Leitão *et al.*, 2025). Furthermore, understanding of the genetic profiles associated with the disease is still limited, highlighting the importance of studies that integrate molecular biology and digital pathology in the characterization of prostate tumors (Hesterberg *et al.*, 2021).

Assessment of PSA levels, digital rectal examination, and imaging exams such as magnetic resonance imaging and ultrasound are the main tools used to detect and diagnose PC (Heidelbaugh *et al.*, 2021). However, definitive diagnostic confirmation still relies on histopathological analysis of prostate tissue obtained by biopsy (Tabayoyong *et al.*, 2015; Patasius *et al.*, 2020). The stratification of the risk of progression and metastasis, as well as the severity of the disease, is predominantly based on PSA levels, Gleason score, and clinical staging (Klein *et al.*, 2022).

Despite the usefulness of these prognostic tools, significant limitations persist. A considerable number of patients initially classified as low-risk may have aggressive disease characteristics after RP (Xia *et al.*, 2020). Adding to this is the growing debate about the true value of PSA as a standalone marker, given the controversies regarding its normal reference levels and its low specificity in distinguishing between indolent and clinically significant tumors (Sundaresan *et al.*, 2025).

Given these limitations, the complementary use of specific protein biomarkers has

shown promise in improving tumor characterization and enabling more precise risk stratification in PC (Saeidi *et al.*, 2023; Falahatkar *et al.*, 2021). Several genes and proteins have been extensively investigated for their involvement in prostate carcinogenesis, particularly those related to tumor suppression, Deoxyribonucleic Acid (DNA) damage response, and immune modulation, such as Phosphatase and Tensin Homolog (PTEN), KI-67 Antigen (KI-67), Ataxia Telangiectasia Mutated (ATM), and the presence of Tumor-Infiltrating Lymphocyte (TIL) (Vilaca *et al.*, 2023; Spyratou *et al.*, 2023; Walker *et al.*, 2021). Among the morphological features with growing prognostic relevance, the cribriform architecture has been particularly associated with adverse outcomes, including higher-grade disease and extraprostatic extension.

Therefore, the present study aims to apply digital pathology tools to investigate the prognostic value of protein biomarkers such as PTEN, Ki-67, ATM, and $CD8^+$ T lymphocytes (CD8) cell density, as well as key histological features including GP4 percentage and CP. In addition, it explores whether digital quantification of PTEN expression can stratify cases beyond traditional binary assessment. Finally, the study examines whether these molecular and histological variables, individually or in combination, are associated with adverse tumor characteristics and whether they can predict the presence of CP in PC.

4.2 Materials and Methods

4.2.1 Sample Collection and Clinical Data Acquisition

Eligibility criteria were defined to ensure sample quality and consistency. Inclusion criteria comprised patients with histopathologically confirmed prostate cancer who underwent radical prostatectomy between 2017 and 2022 and were assisted by the partner laboratory participating in this study. Exclusion criteria included cases with insufficient biological material for the planned analyses, absence of complete clinical or pathological follow-up data, or availability of tissue exclusively from post-neoadjuvant chemotherapy or radiotherapy specimens (biopsies or surgical material).

A total of 200 tumor tissue samples from radical prostatectomies performed between 2017 and 2022 were initially recovered from a reference pathology laboratory. Following histopathological reevaluation, a subset of 72 samples with complete clinical and digital pathology data was included in the final analysis. The samples were stratified into two groups: (1) prostatic adenocarcinoma with a cribriform growth pattern and (2) prostatic adenocarcinoma

without a CP. Relevant clinical data were collected from previous pathology reports and patient medical records. All procedures involving human tissue samples were conducted in accordance with the ethical standards of the institutional and national research committees. This study was approved by the Research Ethics Committee of the Universidade Christus under protocol number CAAE: 67719222.0.0000.5049.

The samples were classified into cribriform and non-cribriform groups based on the morphological evaluation of Hematoxylin and Eosin (H&E) slides by experienced uropathologists. CP was defined according to the contemporary consensus criteria established by the International Society of Urological Pathology (ISUP).

4.2.2 Tissue Microarray (TMA) Construction

TMA were constructed by extracting two 2-mm diameter cores from the most representative tumor areas of RP specimens. These cores were embedded into recipient paraffin blocks arranged in a 6×10 matrix format. To ensure analytical consistency and serve as internal controls, two additional cores from previously validated samples (one positive and one negative control) were included. Two blank cores were also placed at the end of the TMA block to assist with orientation during subsequent analyzes. Sequential 4-µm sections were cut from the TMA block, stained with hematoxylin and eosin H&E for diagnostic confirmation, and subsequently used for immunohistochemical assays.

4.2.3 Immunohistochemistry Analysis

Immunohistochemical analysis was performed using the Dako PT Link and Dako Autostainer 48S Link platforms (CA, USA). Antigen unmasking was carried out by high-temperature heating in a 10 mM citric acid solution (pH 6.0) at 121 ° C for 3 minutes under 25 mmHg pressure. Endogenous erythrocyte peroxidase activity was blocked with 3% hydrogen peroxide (10 volume solution), and nonspecific protein binding was reduced using an antibody diluent. The slides were incubated with the respective primary antibodies for 24 hours at 4 ° C and then further incubated for 1 hour with a secondary antibody. Signal detection was performed using Diaminobenzidine Chromogen (DAB) (Sigma-Aldrich Chemie, Steinheim, Germany), and the slides were counterstained with Harris hematoxylin (Merck) prior to cover slipping and light microscopy analysis.

4.2.4 Image Acquisition and Digital Analysis with QuPath

H&E- and immunohistochemistry-stained slides from the three TMAs were digitized at 40x magnification using the MoticEasyScan platform. Tumor regions were manually annotated by a pathologist on the H&E-stained slides and used as reference for digital analysis of the immunostained sections in QuPath (https://qupath.github.io).

Cell detection was performed using the Cell Detection and Positive Cell Detection commands, based on the optical density of the hematoxylin staining and adjusted positivity parameters. A random tree classifier was interactively trained with manual annotations from a pathologist to distinguish tumor cells from stroma, as well as to categorize tumors as low-grade or high-grade.

The expression of PTEN, ATM, and Ki-67 was assessed using the trained algorithm, considering only tumor cells, by extracting the mean cytoplasmic positivity (DAB stain intensity). Specifically for PTEN and ATM, the H score (0–300) was calculated according to the methodology described by Bankhead et al. (2017).

The density of CD8+ cell was estimated by tissue segmentation (simple tissue detection), followed by an automated count of positive cells, expressed as cells/mm².

4.2.5 Statistical Analysis

All statistical analyzes were performed with RStudio (R version 4.3.2; R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were summarized using means, medians, standard deviations, and interquartile ranges, depending on the distribution. Associations between categorical variables were evaluated using the chi-square test or Fisher's exact test, as appropriate. Comparisons between categorical and continuous variables were made using the Mann-Whitney U test or the Kruskal-Wallis test. Correlations between continuous variables were assessed using the Spearman rank correlation coefficient.

To explore associations with CP, a multivariate logistic regression model was constructed using continuous variables, and the results were expressed as Odds Ratio (OR) with 95% confidence intervals. The predictive performance of the final model was evaluated using ROC curve analysis and the AUC was calculated to assess discriminative ability. Cutoff values derived from the ROC analysis were used to generate categorical variables for some markers (e.g., PTEN H score).

All statistical tests were two-tailed, and significance was set at $p < 0.05$ for all comparisons.

4.3 Results

4.3.1 Sample Collection and Clinical Data Acquisition

The mean age of the 72 patients in the cohort was 66.82 years, with a standard deviation of 7.39, ranging from 50 to 84 years. The clinicopathological characteristics of the cohort were evaluated using data from previous pathology reports and medical records (Table 2).

Table 2 – Clinicopathological characteristics of the patients in the cohort.

	n = 72 (%)
Gleason Score	
6	2 (2,8)
7	57 (79,2)
8	4 (5,6)
9	9 (12,5)
Cribriform Histology	
Yes	37 (51,4)
No	35 (48,6)
Prognostic Group	
1	3 (4,2)
2	34 (47,2)
3	22 (30,6)
4	6 (8,3)
5	7 (9,7)
Pathologic Stage	
T2	33 (47,8)
T3	36 (52,2)
Margins	
Positive	25 (34,7)
Negative	44 (61,1)
Organ-confined tumor	
Yes	31 (43,1)
No	37 (51,4)
Lymphovascular Invasion	
Present	1 (1,4)
Absent	69 (95,8)

Source: prepared by the author (2025).
Note: Percentages may not total 100% due to rounding.

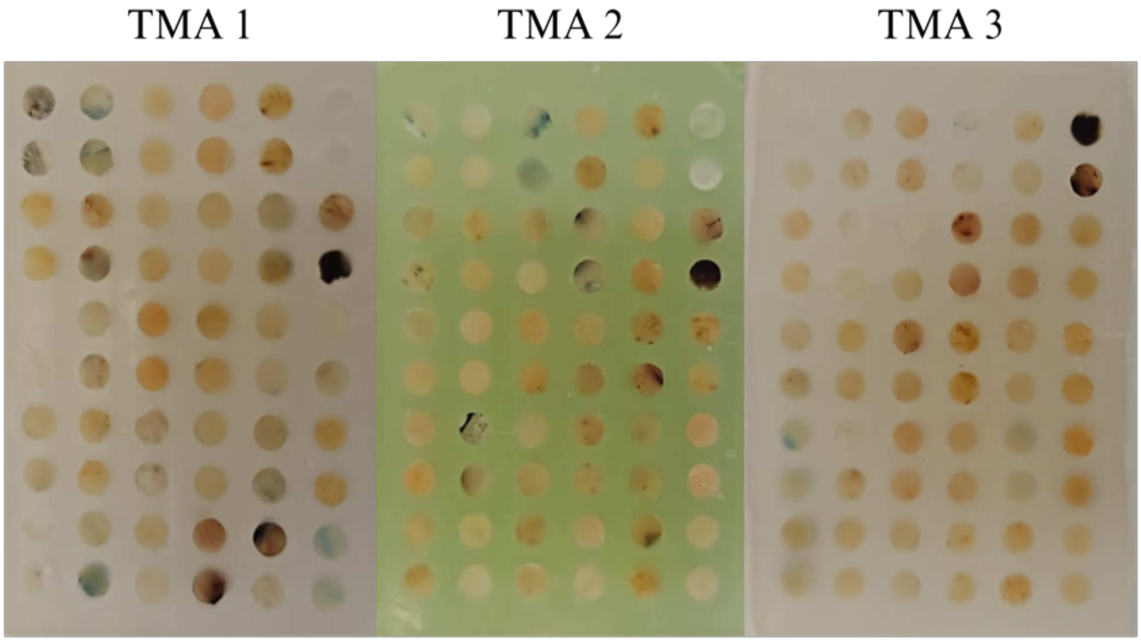
Most samples had a global Gleason score of 7 (79.2%), with half of them presenting a CP (51.4%). Regarding the prognostic group, the most frequent samples were classified as group 2 (47.2%) and group 3 (30.6%). Most samples had a pathological stage of T2 (43.1%), with

extracapsular extension (51.4%) and tumors not confined to the prostate (51.4%). However, most patients also had negative surgical margins (61.1%) and no lymphovascular invasion (95.8%).

4.3.2 Tissue Microarray constrution (TMA)

A total of 86 Formalin-Fixed Paraffin-Embedded (FFPE) tumor tissue samples from patients diagnosed with PC were selected for the construction of the TMA. Three paraffin blocks were built, each containing 50 donor sample cores of 2 mm in diameter (Figure 8). After constructing the three TMAs, they were subjected to histological sectioning for the preparation of slides for staining with H&E and immunohistochemical.

Figure 8 – TMA constructed from representative tumor cores of radical prostatectomy specimens included in the study cohort.

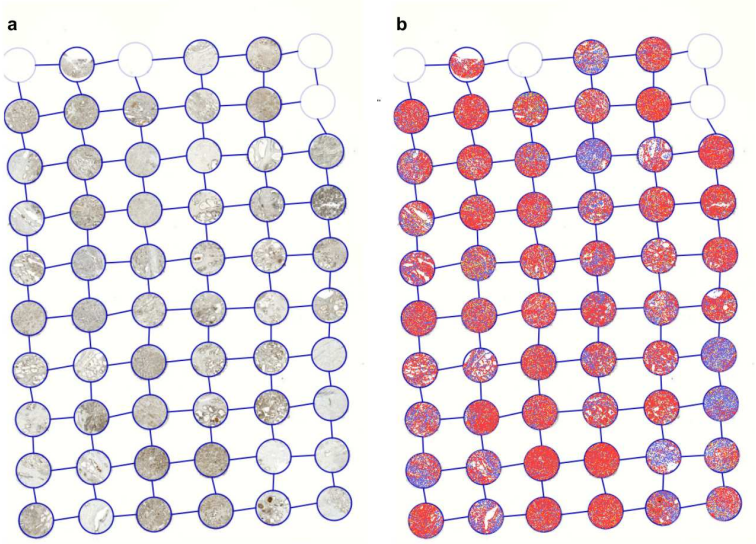


Source: prepared by the author (2025).

4.3.3 Immunohistochemistry and image analysis

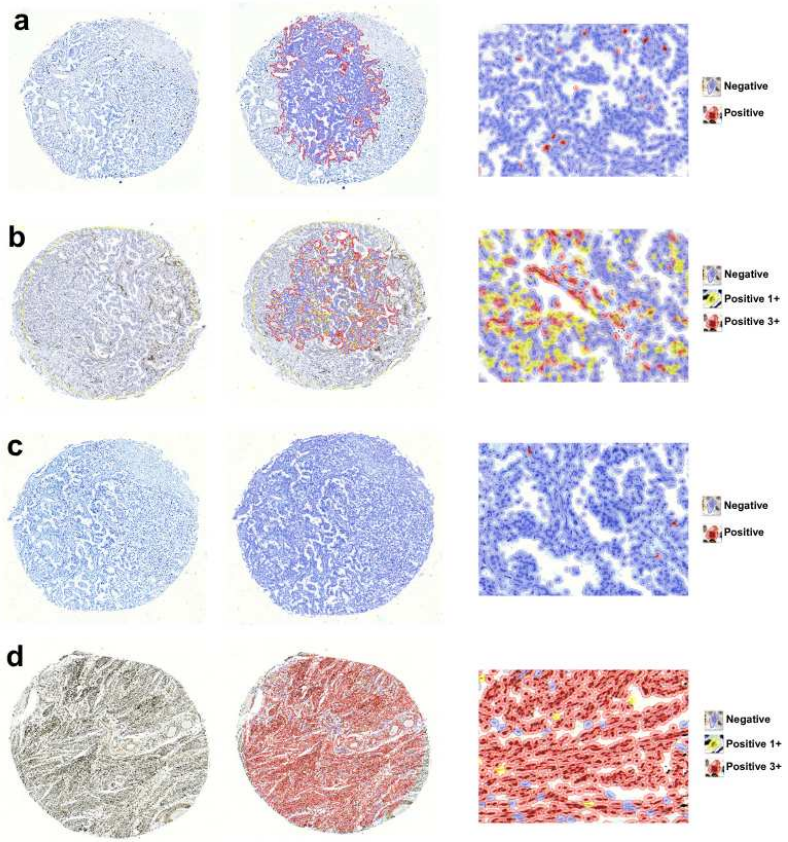
The H&E slides from the three TMAs were scanned at 40x using the MoticEasyScan platform. With the assistance of an experienced pathologist, the tumor regions of each sample (cores) were marked. Subsequently, the immunohistochemical reaction slides were digitized and analyzed using QuPath software, employing tumor markings (Figures 9 and 10).

Figure 9 – Tissue microarray (TMA) samples of prostate cancer (PC). (a) Hematoxylin and eosin (H&E) staining of TMA cores. (b) Automated detection of nuclear ATM protein expression by immunohistochemistry, performed using QuPath software.



Source: prepared by the author (2025).

Figure 10 – Automated quantification of immunohistochemical staining for (a) PTEN, (b) CD8, (c) Ki-67, and (d) ATM on tissue microarray (TMA) cores using QuPath. The software identified positive and negative areas and, when applicable, graded staining intensity (1+, 3+) as shown in the color-coded maps.

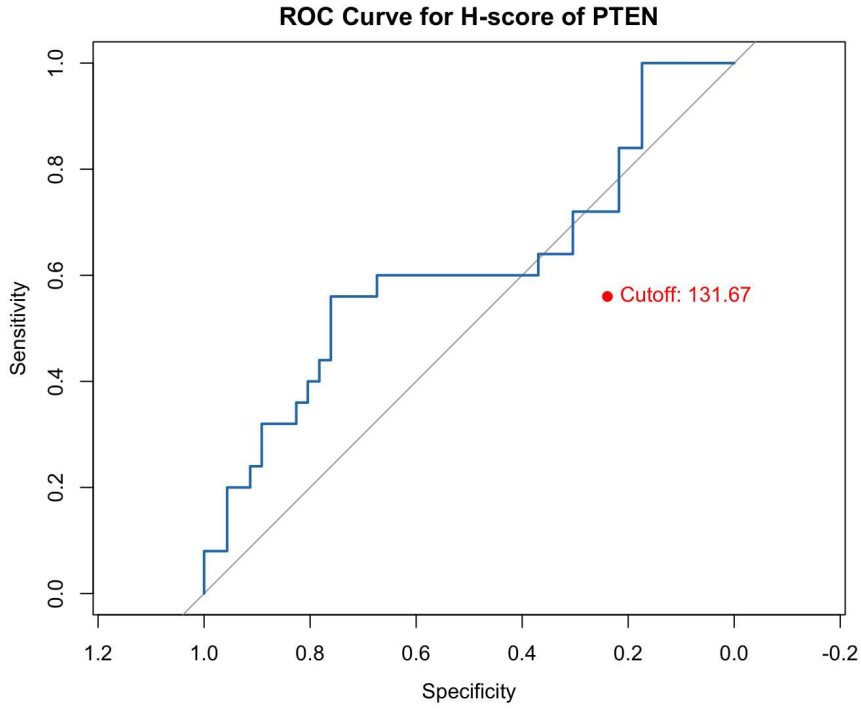


Source: prepared by the author (2025).

4.3.4 Prognostic Value of PTEN H-score Quantification via Digital Pathology

PTEN is one of the most studied tumor suppressor proteins in PC, with loss of PTEN associated with a poor prognosis and aggressive disease phenotypes. Although traditional pathology often reports PTEN as "intact" or "lost", digital pathology enables a continuous quantification of PTEN expression through H-score, potentially allowing for a more refined prognostic stratification. Descriptive analysis of PTEN H-score values revealed a mean of 158.56 and a median of 150.31, with values ranging from 50.66 to 266.21. The difference between the mean and median suggests a slight right-skewed distribution, indicating the presence of cases with higher protein expression. The interquartile range was approximately 101.04 (1st quartile = 113.90; 3rd quartile = 214.94), reflecting substantial variability in expression levels among the cases analyzed. These findings demonstrate a broad distribution of PTEN H scores in the studied cohort, which may reflect the biological heterogeneity of prostate tumors. To assess the ability of the PTEN H score to discriminate cases with protein loss (as evaluated by the pathologist), a ROC curve analysis was performed(Figure 11).

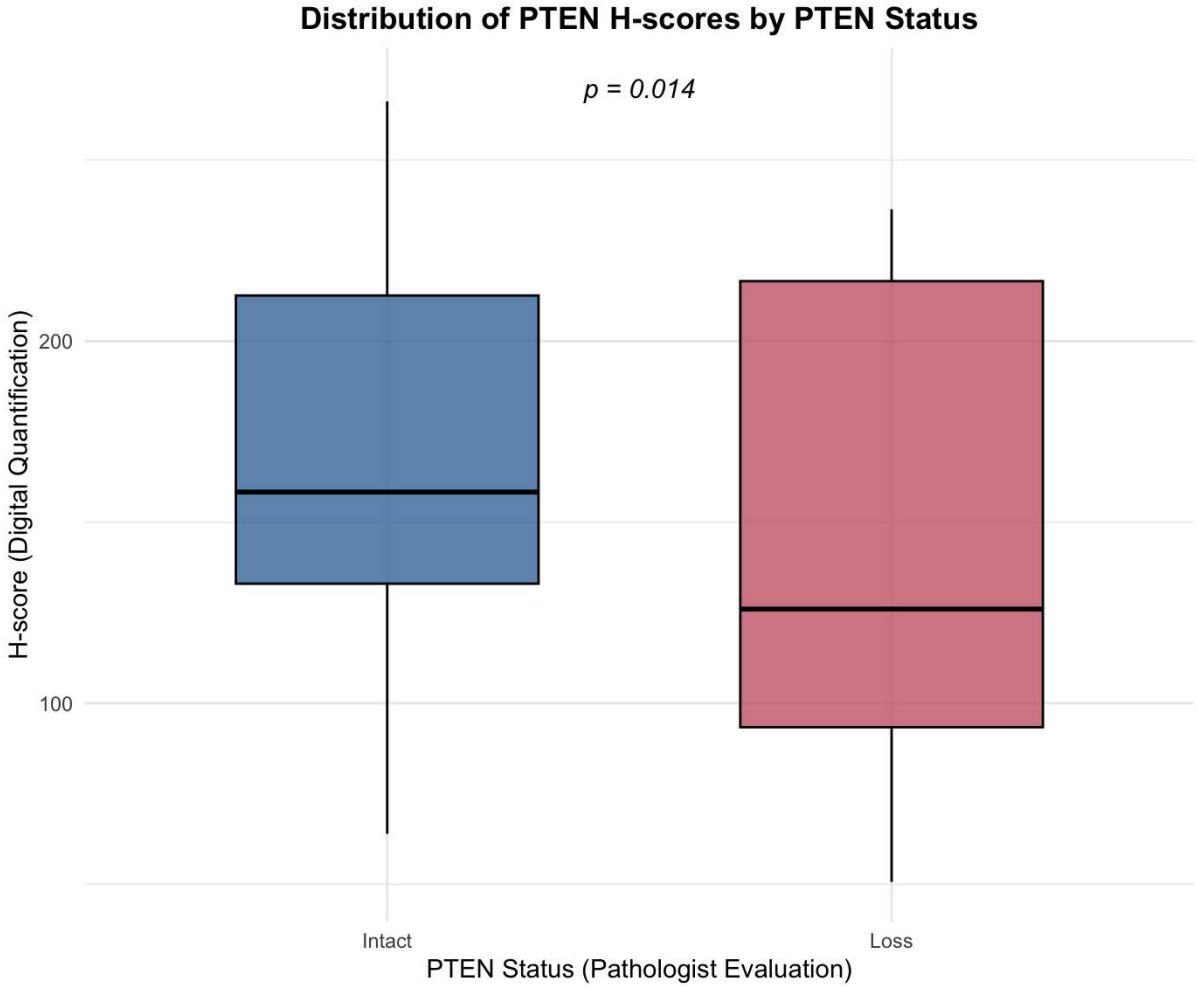
Figure 11 – Receiver operating characteristic (ROC) curve analysis of PTEN H-score for discriminating cases with protein loss. The area under the curve (AUC) was 0.610, with an optimal cut-off point of 131.67, yielding a sensitivity of 56% and a specificity of 76%



Source: prepared by the author (2025).

The AUC was 0.610, indicating a low discriminatory capacity between the groups. The identified optimal cut-off point was 131.67, yielding a sensitivity of 56% and a specificity of 76%. Based on this threshold, cases were classified as having low or high expression of PTEN, allowing an objective classification of patients using digital quantification. This new categorical variable showed a statistically significant association with the binary PTEN status evaluated by the pathologist (Chi-square test, $p = 0.014$(Figure 12)).

Figure 12 – Distribution of PTEN H-score according to PTEN protein status evaluated by the pathologist. Boxplot showing the digital quantification of PTEN (H-score) stratified by cases with intact PTEN expression and those with PTEN loss (binary evaluation by immunohistochemistry). A statistically significant association was observed between the categorical H-score variable (based on the ROC-derived cutoff) and the binary PTEN status (Chi-square test, $p = 0.014$).

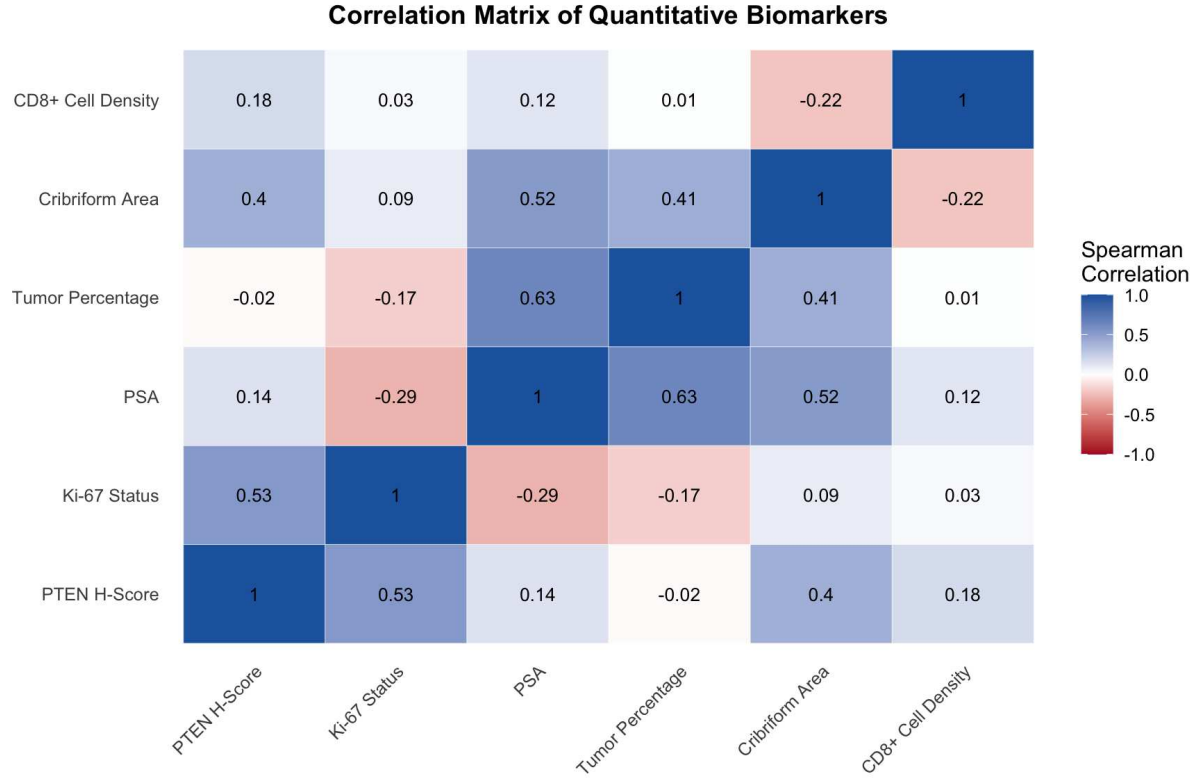


Source: prepared by the author (2025).

4.3.5 Integrated Analysis of Quantitative Markers Highlights Tumor Aggressiveness Profile in Prostate Cancer

The correlation matrix of continuous quantitative variables (Figure 13) revealed distinct patterns among biomarkers related to tumor aggressiveness. Tumor percentage, PSA, cribriform area, and Ki-67 status showed moderate to strong positive correlations (r = 0.4–0.63), indicating a cluster of variables associated with higher tumor burden and proliferative activity. CP was positively correlated with PSA (r = 0.52) and tumor percentage (r = 0.41), reinforcing its association with adverse pathological features. PSA also demonstrated a relevant correlation with tumor percentage (r = 0.63), further linking serum levels to tumor extent. PTEN H-score showed a moderate positive correlation with Ki-67 status (r = 0.53) and CP (r = 0.40). This unexpected relationship will be further addressed in the discussion. CD8+ cell density did not exhibit relevant correlations with the other biomarkers (r between -0.22 and 0.18), reflecting a tumor microenvironment with limited cytotoxic immune infiltration.

Figure 13 – Correlation matrix of quantitative variables in PC. Strong correlations were observed among Ki-67, GP4, tumor extent, PTEN loss, and CP. CD8+ cell density and PSA showed weak or no correlations.

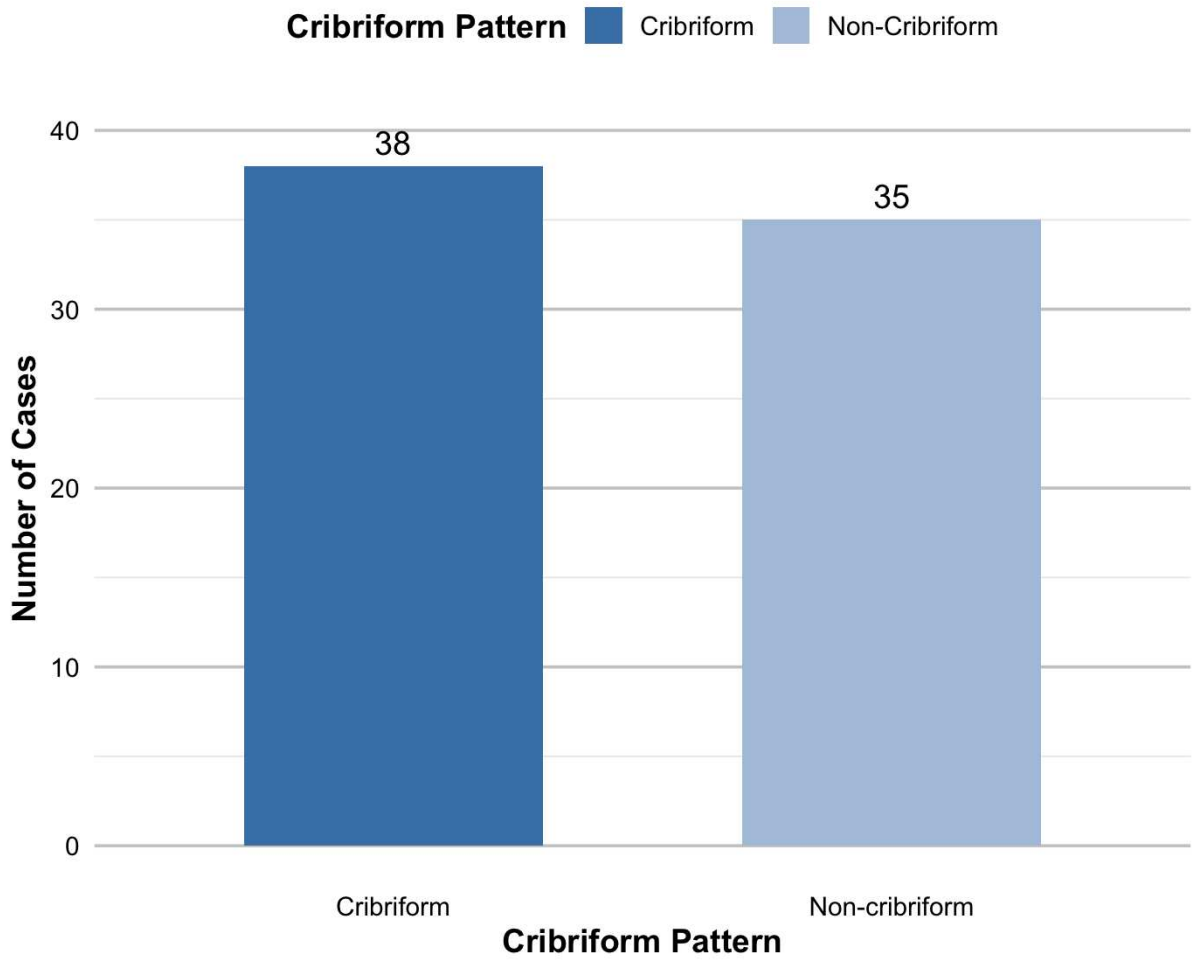


Source: prepared by the author (2025).

4.3.6 Association of Cribriform Pattern with Adverse Pathological Features

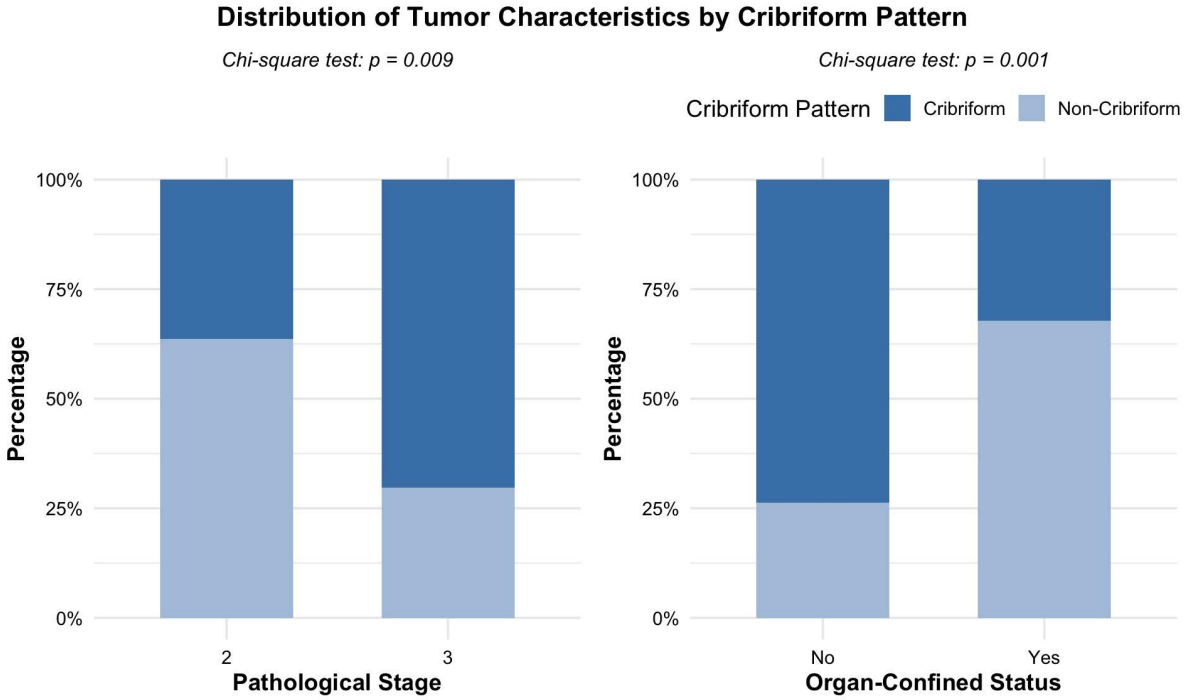
The CP classification was performed by experienced pathologists, resulting in 38 samples classified as cribriform and 35 as non-cribriform. As shown in Figure 14, the two groups were relatively balanced in number. A statistically significant association was found between the presence of the CP and the pathological stage ($p = 0.009$), suggesting that cribriform tumors are diagnosed more frequently in advanced stages. Furthermore, cribriform morphological tumors were significantly less likely to be organ-confined ($p = 0.001$), indicating a more aggressive clinical behavior. These associations are illustrated in Figure 15, which highlights the prognostic value of the assessment of the cribriform morphology in PC.

Figure 14 – Bar plot showing the number of PC cases categorized as cribriform (n = 38) and non-cribriform (n = 35), based on the morphological evaluation performed by experienced pathologists.



Source: prepared by the author (2025).

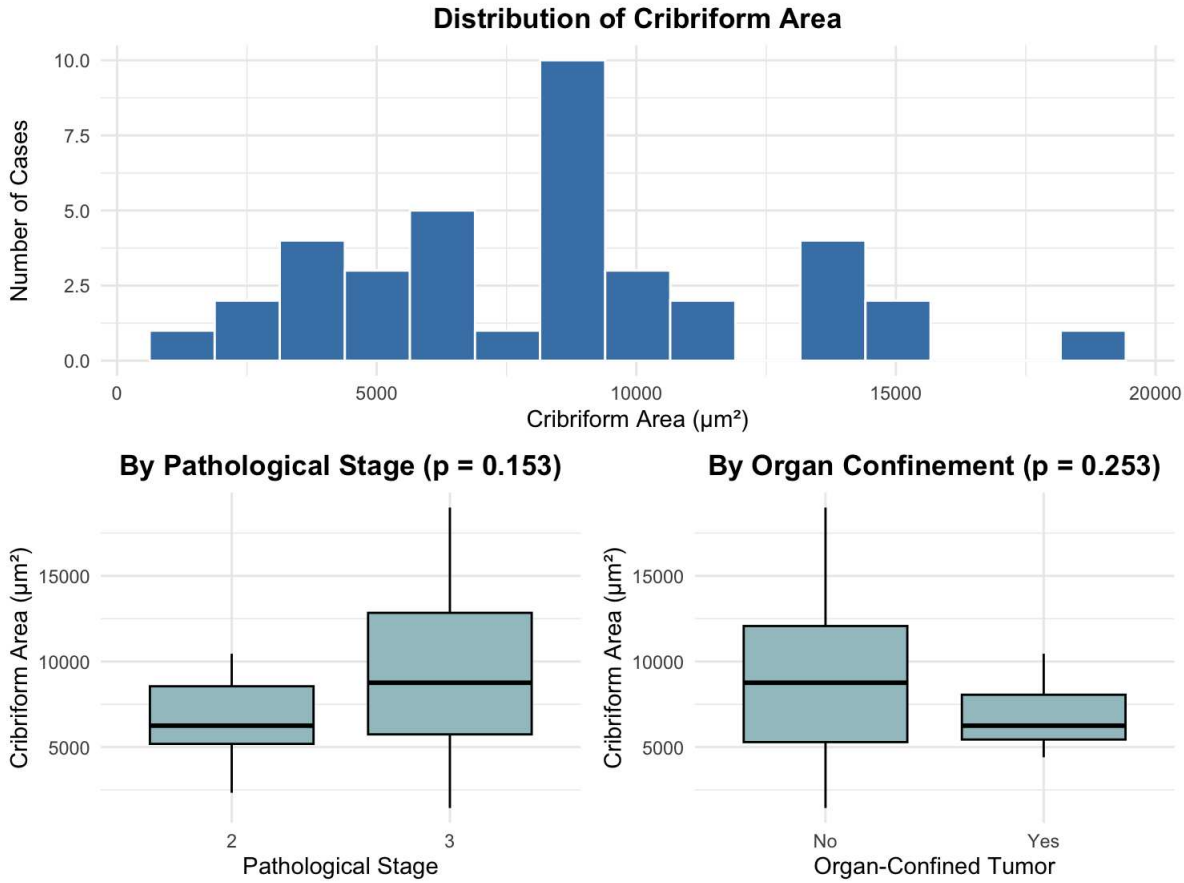
Figure 15 – Proportion of cribriform and non-cribriform cases across (A) pathological stage (Chi-square test, p = 0.009) and (B) organ-confined status (Chi-square test, p = 0.001). The presence of CP was significantly associated with higher-stage tumors and lower frequency of organ-confined disease.



Source: prepared by the author (2025).

The analysis of the cribriform area, quantified through digital pathology, revealed a wide range of values in the samples, with a mean of 8,348 µm² and a median of 8,532 µm² (Figure 16, top). This distribution reflects the biological heterogeneity of the CP observed in prostate tumors. When stratified by pathological stage, cases with stage III disease tended to have slightly larger cribriform areas compared to those with stage II; however, this difference did not reach statistical significance (Wilcoxon test, p = 0.153) (Figure 16, bottom left). Similarly, no significant differences were observed in the cribriform area between tumors confined to the organ and those with extraprostatic extension (p = 0.253) (Figure 16, bottom right). These findings suggest that, although the presence of CP is associated with a worse prognostic feature, its quantitative extent may not independently discriminate between pathological stages or tumor confinement status.

Figure 16 – Cribriform area quantification in prostate tumors. Top: Histogram of cribriform area (µm²). Bottom: Boxplots showing comparisons by pathological stage (left, $p = 0.153$) and organ confinement (right, $p = 0.253$). No significant differences observed.

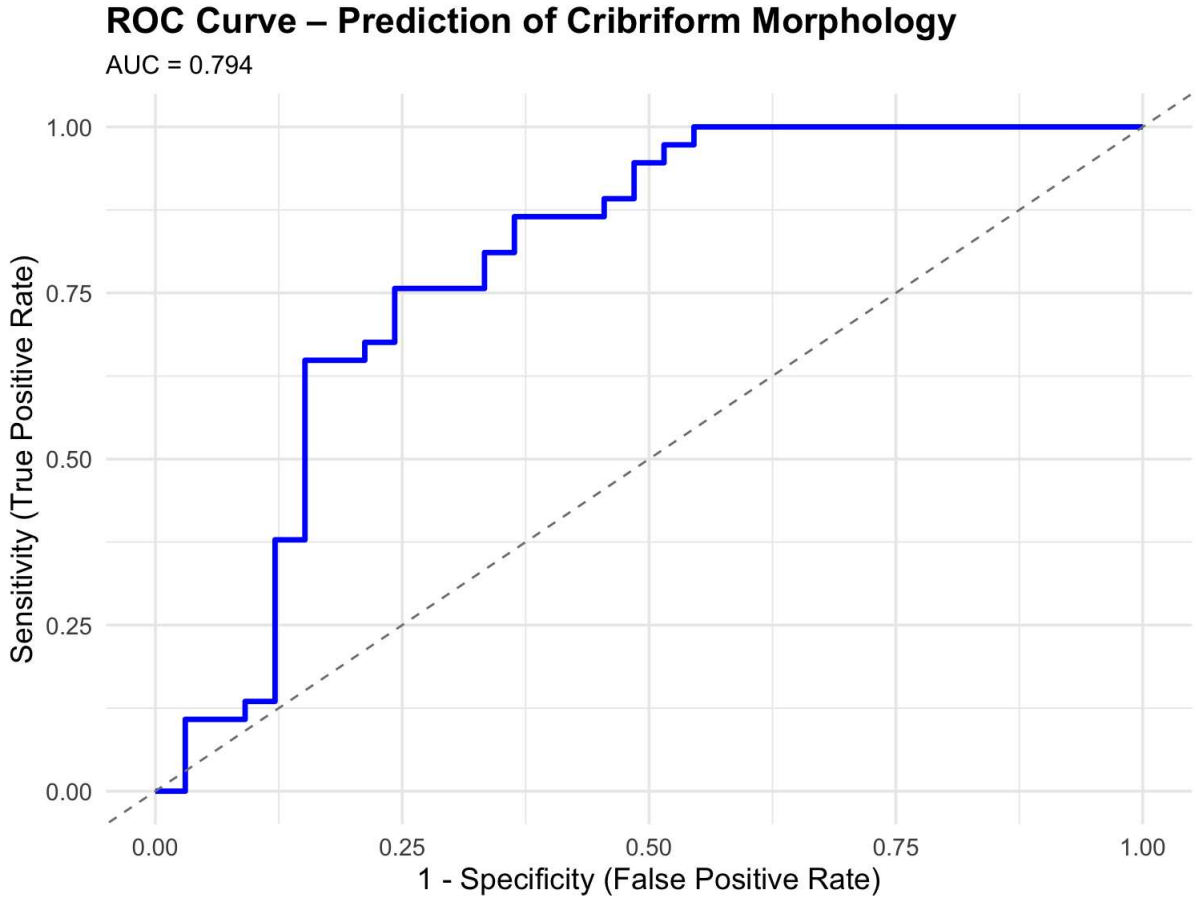


Source: prepared by the author (2025).

4.3.7 Multivariate Analysis and ROC Performance for Cribriform Pattern Prediction

A multivariate logistic regression model was fitted that included the percentage of GP4 and ATM protein expression (H-score) was fitted to assess their association with CP in PC samples (Figure 17). Both variables were significantly associated with the outcome. Specifically, each 1% increase in GP4 was associated with a 4.2% increase in the odds of having a CP (OR = 1.04; 95% Confidence Interval (CI): 1.02–1.07; $p < 0.001$). ATM expression also demonstrated a significant association, with a 1.8% increase in odds for each unit increase in the h score (OR = 1.02; 95% CI: 1.00–1.04; $p = 0.049$). The discriminative performance of the model was evaluated using a ROC curve, resulting in an AUC of 0.79, indicating a good predictive ability.

Figure 17 – ROC curve for the logistic regression model predicting CP based on GP4 percentage and ATM expression. The model showed good discriminative performance with an AUC of 0.79.



Source: prepared by the author (2025).

4.4 Discussion

The clinicopathological profile of the cohort was consistent with typical PC populations reported in the literature, with a predominance of intermediate-risk cases (Gleason score 7 and prognostic groups 2 and 3) and a mean age of approximately 67 years (Ham *et al.*, 2017). The balanced distribution of cribriform and non-cribriform cases (51.4% vs. 48.6%) provided an appropriate basis for comparative analyses. Despite a considerable proportion of cases showing extraprostatic extension and non-organ-confined disease, the high frequency of negative surgical margins and the absence of lymphovascular invasion suggest effective surgical management and reinforce the heterogeneity of PC progression patterns (Karwacki *et al.*, 2024).

The use of TMA enabled efficient and standardized analysis of multiple prostate tumor samples within a controlled and reproducible framework. By selecting representative tumor regions and incorporating internal controls, TMAs provided a reliable platform for histological

evaluation and immunohistochemical quantification. This approach is particularly valuable in biomarker studies, as it reduces variability between slides, conserves reagents, and ensures consistent processing between samples (Vlajnic *et al.*, 2018; Ye *et al.*, 2020; Stopsack *et al.*, 2021).

The integration of digital pathology through high-resolution slide scanning and manual tumor annotation by a pathologist ensured the precision of region-of-interest selection for downstream analysis. This combination of expert-driven marking with computational processing in QuPath enabled consistent identification of tumor areas and minimized subjectivity (Bankhead *et al.*, 2017; Greene *et al.*, 2021). The digital workflow also facilitated the scalable quantification of immunohistochemical markers, supporting both reproducibility and data traceability throughout the study (Moratin *et al.*, 2021).

Digital quantification of PTEN expression using H-score provided a continuous and objective measure that captured the heterogeneity of protein expression between tumor samples. The wide range and distribution of H-scores observed in this study reflect the biological variability of PTEN levels in PC, supporting the use of digital tools to move beyond binary classifications (Albuquerque *et al.*, 2019; Hommerding *et al.*, 2024). Although the discriminatory power of the H-score to predict the status of PTEN as assessed by the pathologist was modest (AUC = 0.61), the statistically significant association between the categorical variable derived from ROC and the binary immunohistochemical classification reinforces the potential of digital quantification to complement traditional evaluations. These findings suggest that, while the digital H-score alone may not fully replace expert interpretation, it can enhance reproducibility and contribute to more refined prognostic models when integrated with additional molecular and histological data (Moratin *et al.*, 2021).

The correlation analysis revealed a consistent pattern linking high tumor percentage, increased GP4, elevated Ki-67 expression, and PTEN expression—features commonly associated with biologically aggressive prostate tumors (Dallang *et al.*, 2025; Jamaspishvili *et al.*, 2018; Sharma *et al.*, 2018). This clustering supports the concept of an integrated aggressiveness profile, where proliferative activity, architectural disorganization, and tumor suppressor pathway alterations converge in high-risk cases (Bogaard *et al.*, 2023). The positive correlation between CP and these variables further reinforces its value as a histological indicator of poor prognosis. Notably, the absence of significant correlations involving CD8+ cell density suggests that, within this cohort, more aggressive tumors may exhibit an immunologically “cold” phenotype, lacking

effective cytotoxic immune infiltration (Kwon *et al.*, 2025). These findings highlight the complex interplay between tumor-intrinsic features and the immune microenvironment in PC (Brea *et al.*, 2025).

The significant association between CP and adverse pathological features—such as higher pathological stage and lack of organ confinement—reinforces its role as a marker of aggressive tumor behavior. These findings are consistent with the broader notion that cribriform architecture is not merely a histological variant, but rather a structural manifestation of biologically advanced disease (Shimodaira *et al.*, 2024). Its association with both local invasion and pathological upstaging supports its use in risk stratification and prognostic assessment in PC (Osiecki *et al.*, 2023).

Interestingly, the quantitative measurement of cribriform area did not correlate with either pathological stage or tumor confinement, suggesting that the extent of CP may not provide additional prognostic value beyond its presence. Supporting this notion, Chen *et al.* (2021) quantitatively assessed CP in 145 PC specimens and found no significant association between the measured cribriform area and organ confinement or pathological staging ($p > 0.05$). These findings underscore the clinical relevance of identifying CP qualitatively—based on its presence—rather than attempting to quantify its extent, particularly within the context of diagnostic pathology workflows.

The multivariate analysis identified both GP4 percentage and ATM expression as independent predictors of CP, reinforcing the interplay between histological architecture and molecular alterations in PC. The strong association with GP4 was expected, given its well-established link to tumor aggressiveness (Sharma *et al.*, 2018; Kroon *et al.*, 2025). Notably, the inclusion of ATM H-score as a significant predictor highlights a potentially underexplored role for this DNA damage response marker in driving or reflecting cribriform growth (Kaur *et al.*, 2020). While few studies have directly examined the relationship between ATM expression levels and CP, Wong *et al.* (2022) reported that prostate tumors with cribriform architecture exhibit pronounced genomic instability and activation of DNA repair pathways. Given ATM's central role in the DNA damage response, these findings support the biological plausibility that altered ATM expression may contribute to the emergence or maintenance of this aggressive histological pattern. The resulting model demonstrated good discriminative performance (AUC = 0.79), suggesting that the integration of morphological and molecular markers through digital quantification may serve as a valuable approach for predicting high-risk histological patterns. These findings support

the use of targeted immunohistochemical panels to enhance histopathological risk stratification, although further validation in independent cohorts is warranted.

Together, these findings underscore the potential of digital pathology to integrate morphological and molecular features into clinically meaningful insights, offering a scalable and reproducible approach to improve prognostic stratification in PC.

4.5 Conclusion

This study underscores the utility of digital pathology as a powerful platform for the integrated assessment of molecular and morphological features in PC. By quantifying immunohistochemical markers and histological patterns, we identified the proportion of GP4 and ATM expression as independent predictors of CP, a morphology consistently linked to tumor aggressiveness. The presence of cribriform architecture was strongly associated with adverse pathological parameters, reinforcing its prognostic relevance. Although digital PTEN quantification offered valuable insights into tumor heterogeneity, its predictive capacity proved limited when considered alone. Collectively, these findings support the integration of digital pathology with targeted biomarker panels as a strategy to refine risk stratification and guide personalized clinical decision-making in PC. Future studies in independent cohorts are warranted to validate these observations and to advance their translation into routine diagnostic practice.

5 GENE EXPRESSION SIGNATURES IN CRIBRIFORM PROSTATE CANCER: AN *IN SILICO* APPROACH

Francisco Araujo[1], Emerson Lucena[2], Guilherme Velozo[2], Juliana Cordeiro[2], Adriane Feijó[3], Denise Hissa[1], Vania Melo[1], Fábio Távora[1,2]

[1]Department of Biology, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil

[2]Laboratório Argos Patologia, Fortaleza, Ceará, Brazil

[3]Department of Genomic, Instituto de Pesquisa Genômica, Guarapuava, Paraná, Brazil

Submission date: July 17, 2025

Journal submitted: Archives in Pathology

Status of submission: Under review

Abstract

Cribriform morphology in prostate cancer is associated with worse prognosis, including higher risk of recurrence, metastasis, and specific mortality. However, its underlying molecular mechanisms remain poorly understood. This study aimed to characterize the gene expression profile associated with cribriform morphology in prostate cancer samples, using RNA-Seq data from the GEO-NCBI database and applying robust statistical strategies to overcome limitations of sample imbalance. In our analyses, we identified 55 differentially expressed genes between cribriform and non-cribriform samples. Genes such as *CPNE4*, *PLA2G7*, and *PTPRJ* were detected in both analyses, while *GCNT1*, *AMACR*, *AR*, and *ERG* were overexpressed in cribriform samples, being associated with immune evasion, lipid metabolism, and hormonal pathways. *TRIM9*, *HOXD3*, and *ITGB3BP* were underexpressed, suggesting a loss of regulatory functions in the tumor microenvironment. Functional analysis revealed a transcriptional reprogramming associated with tumor aggressiveness, with an emphasis on phosphodiesterase activity and hormonal response. Stratified sampling based on Euclidean distance ensured biological representativity, overcoming sample bias. The results reinforce the association between cribriform morphology and a pro-tumorigenic transcriptional profile, expanding the understanding of molecular pathways involved in aggressive prostate cancer progression.

Keywords: Cribriform morphology, Prostate cancer, Gene expression profile, RNA-Seq analysis.

5.1 Introduction

PC is one of the most common cancers among men, second only to lung cancer. In 2020, according to Globocan, approximately 1.4 million PC cases were diagnosed worldwide, causing around 375,304 deaths among men of all ages (Sung *et al.*, 2021). The main identified risk factors for PC include age, family history, and ethnicity. Other elements such as obesity, unbalanced diets, tobacco, and alcohol consumption have also been linked to PC development (Perdana *et al.*, 2016).

PC is known to be an aggressive and heterogeneous disease with diverse genetic, molecular, and epidemiological profiles, resulting in highly variable outcomes among patients (Sekhoacha *et al.*, 2022). Currently, available risk stratification primarily relies on the Gleason score (Bogaard *et al.*, 2024). New molecular tools, such as biomarkers, are needed to improve diagnosis and guide clinical decision-making.

Given the high heterogeneity of PC, one widely used strategy in disease research is the evaluation of the prognostic value of specific morphological features, such as the CP (Dong *et al.*, 2013; Bogaard *et al.*, 2024). This pattern, a subtype of GP4, is characterized by clusters of tumor cells with intercellular lumens and may be found in both Invasive Cribriform Carcinoma (ICC) and Intraductal Carcinoma (IDC) (Russo *et al.*, 2023). Several studies have shown a strong association between the presence of CP and unfavorable clinical outcomes (Kweldam *et al.*, 2015; Hollemans *et al.*, 2019; Miura *et al.*, 2020). Men with CP at RP are more likely to experience biochemical recurrence, metastatic progression, and disease-specific mortality, regardless of Gleason score.

Genomic studies in samples with CP have revealed multiple molecular aberrations, including genetic instability, copy number alterations, and mutations in genes such as *Ataxia telangiectasia mutated (ATM), Speckle type BTB/POZ protein (SPOP), Breast cancer type 2 susceptibility protein (BRCA2), Tumor protein p53 (TP53), RB transcriptional corepressor 1 (RB1), Marker of proliferation Ki-67 (MKI67), and Phosphatase and tensin homolog (PTEN).* (Elfandy *et al.*, 2019; Lozano *et al.*, 2021; Bottcher *et al.*, 2018). Additionally, these tumors demonstrate a downregulation of apoptotic pathways alongside an upregulation of proliferative and MYC signaling pathways (Lone *et al.*, 2022).

In summary, literature findings support the association between cribriform architecture and worse prognosis. However, due to the molecular heterogeneity of PC, further studies using new cohorts are necessary to identify novel biomarkers capable of improving diagno-

sis and prognostication, and to better understand the molecular mechanisms that underlie the aggressiveness of cribriform tumors.

In this chapter, we performed a Differential Expression Analysis (DEA) using publicly available RNA Sequencing (RNA-seq) datasets, employing rigorous statistical approaches to address sample imbalance issues found in previous studies and to explore the molecular drivers of cribriform tumor aggressiveness.

5.2 Materials and Methods

5.2.1 Data source

The main methodological steps followed in this study are summarized in Figure 18. The human expression profiles of PC were obtained from GEO-NCBI (https://www.ncbi.nlm.nih.gov/gds), which were derived from research carried out by Ma *et al.* 2023. The raw Salmon profiles with the accession number GSE214940 were obtained to identify DEGs in cribriform and non-cribriform PC. The samples analyzed in this study consist of 272 tumor-adjacent and 120 benign-adjacent macrodissected prostate stromal samples from 293 men enrolled in the Health Professionals Follow-up Study and the Physicians' Health Study.

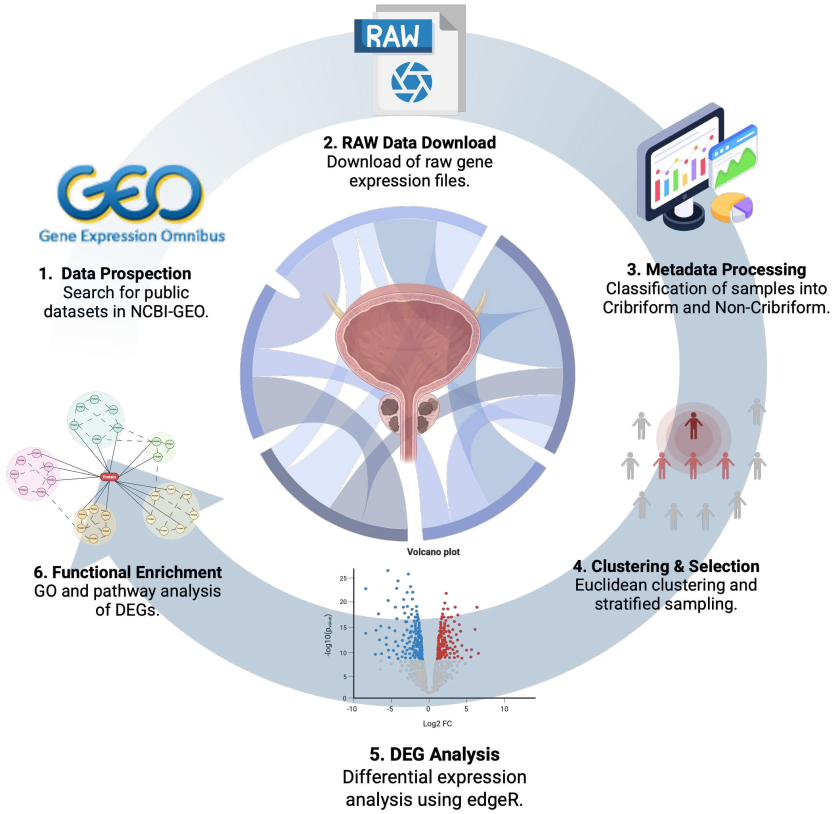
5.2.2 Stratified Control Sample Selection

To minimize the imbalance between groups in the DEA, we implemented a clustering and stratified selection method. The Non-Cribriform group samples (n = 205) were clustered based on Euclidean distance, and the optimal number of clusters was determined using the elbow method (Bholowalia *et al.* 2014). A stratified weighted selection was then performed to obtain a subset of 65 samples from the Non-Cribriform group, matching the size of the Cribriform group (n = 65). The selection was conducted randomly within each cluster, proportionally to its original size, ensuring representativity. This process was automated using an R script, ensuring reproducibility and avoiding biases.

5.2.3 Differentially expressed Genes (DEGs)

Gene counts were computed by use of R (v.4.1.1) (https://www.r-project.org/) and Rstudio (2024.04.2+764) (https://posit.co). DEA was performed using edgeR (v.3.34.1)

Figure 18 – Schematic overview of the methodological workflow, from data prospection to functional enrichment analyses.



Source: prepared by the author (2025).

(https://bioconductor.org/packages/release/bioc/html/ edgeR.html). Count normalization was performed using the trimmed mean M-value method in edgeR. To identify a significant difference in expression an adjusted p-value threshold of 0.05 were used. The results we plotted using the ggplot2 package (v.3.4.2) (https://ggplot2.tidyverse.org).

5.2.4 Gene Ontology and Pathway Analysis

To explore the functional significance of the upregulated genes in the Cribriform group, we performed a functional enrichment analysis using Enrichr (Kuleshov *et al.* 2016). This tool allows for the identification of enriched biological pathways and molecular functions. Specifically, we focused on two categories: MF and BP, within the GO framework. We first selected the upregulated genes from DEA in the Cribriform group and inputted them into Enrichr to identify overrepresented terms in both MF and BP. The enrichment results were filtered based on the p-value, with a threshold of 0.05, to ensure the statistical relevance of the identified terms. The functional categories that showed significant enrichment were further analyzed to understand their potential biological roles in the context of the study.

5.3 Results

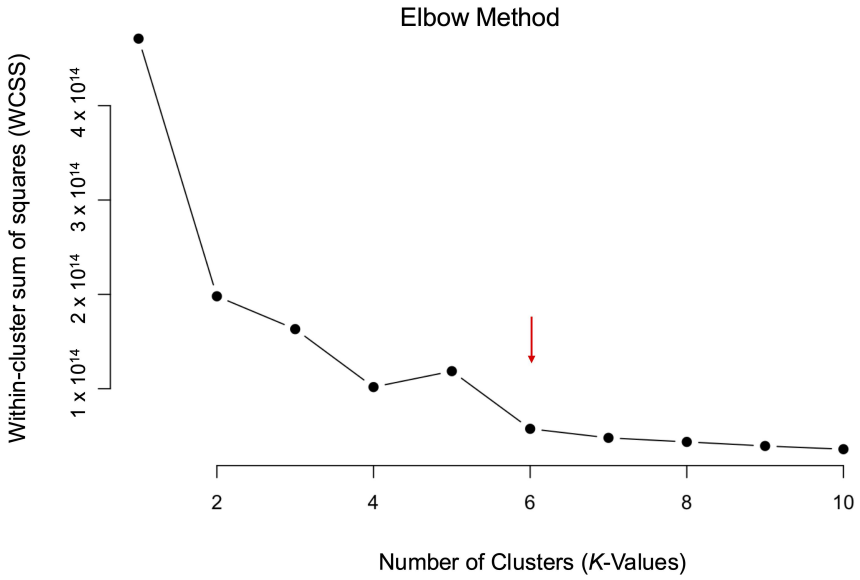
5.3.1 Data classification

After an online screening, only GSE214940 met our research criteria, as it provided RNA-seq quantification data that included the histopathological classification of cribriform and non-cribriform patterns. This dataset comprises RNA-seq data from 293 men, including 272 tumor-adjacent and 120 benign-adjacent prostate stroma samples. Among the 272 tumor-adjacent samples, 205 were classified as non-cribriform, 65 as cribriform, and 2 remained unclassified. These samples were subsequently filtered and used in our analyses.

5.3.2 Stratified Control Sample Selection

To address the imbalance between groups in the DEA, we implemented a stratified control sample selection approach. The Non-Cribriform group initially consisted of 205 samples, while the Cribriform group had 65 samples. Using Euclidean distance, we clustered the Non-Cribriform samples, determining the optimal number of clusters through the elbow method (Figure 19).

Figure 19 – Elbow test result showing the characteristic curvature, indicating the optimal number of clusters.



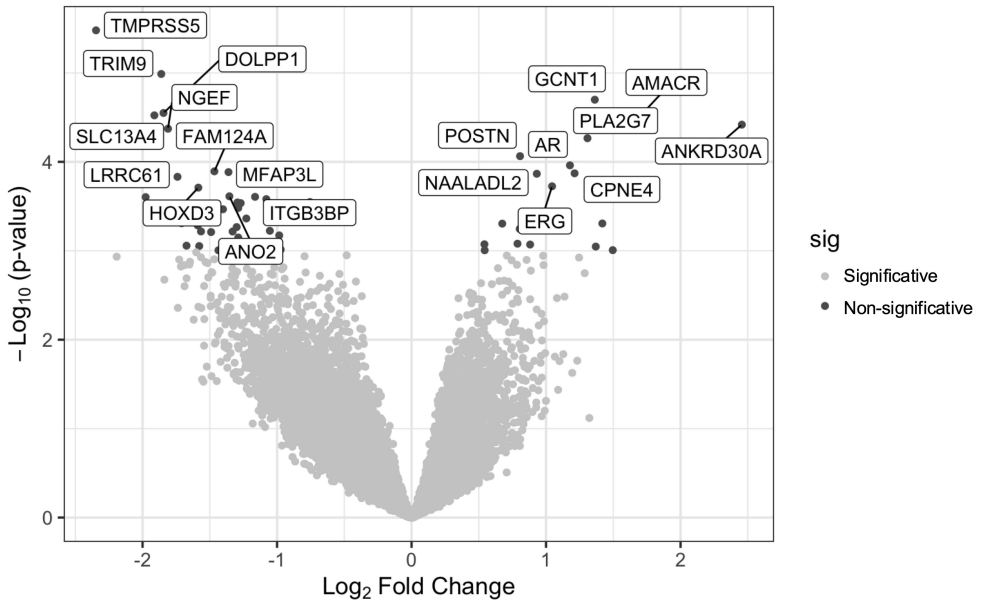
Source: prepared by the author (2025).

Each sample was then assigned to a cluster, and a stratified weighted selection was performed to obtain 65 Non-Cribriform samples, ensuring numerical balance with the Cribriform group. The selection process was automated using an R script, which randomly selected samples from each cluster in proportion to its original size, minimizing selection bias. As a result, we obtained a balanced dataset of 65 Cribriform and 65 Non-Cribriform samples, allowing for a more robust and unbiased DEA.

5.3.3 Differentially expressed Genes (DEGs)

In our analyses, we identified 55 DEGs ($p < 0.01$) between cribriform and non-cribriform PC samples. Among these, 18 genes were upregulated (Log Fold Change (LogFC) $>$ 0) and 37 were downregulated (LogFC $<$ 0).The top 20 genes are highlighted in the volcano plot (Figure 20).

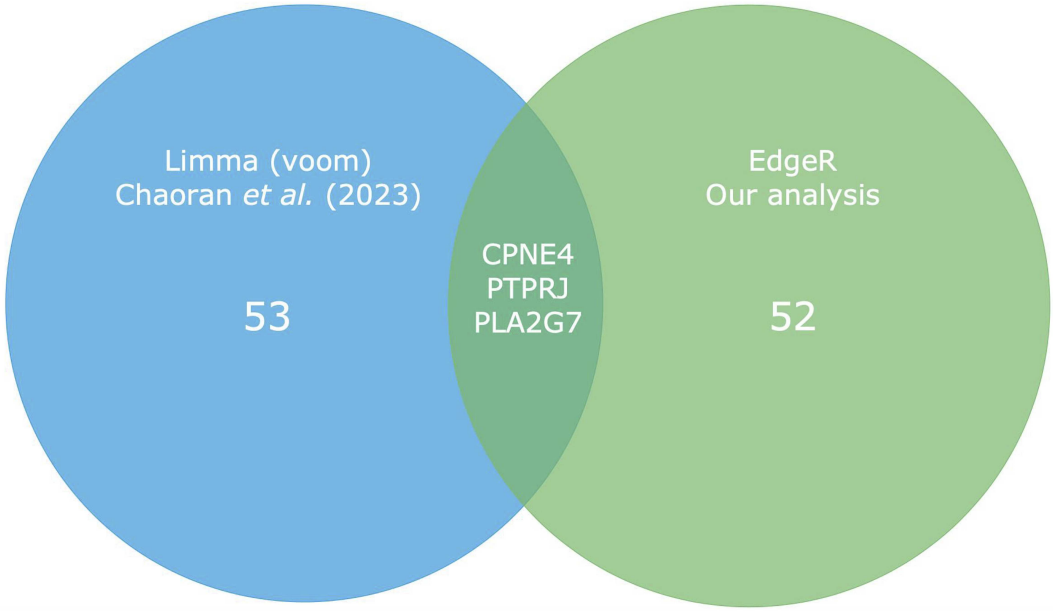
Figure 20 – Volcano plot displaying the top 20 DEGs between cribriform and non-cribriform PC samples.



Source: prepared by the author (2025).

Notably, the genes *Copine 4* (*CPNE4*), *Phospholipase A2 group VII* (*PLA2G7*), and *Protein Tyrosine Phosphatase Receptor Type J* (*PTPRJ*) were detected in both our analysis and the original study, as shown in the Venn diagram (Figure 21).

Figure 21 – Venn diagram comparing the DEGs identified in the original study by Chaoran *et al.* (2023) and our analysis. The diagram highlights the three genes commonly identified in both studies, emphasizing the overlap between the two datasets.



Source: prepared by the author (2025).

Prominent upregulated genes included *Ankyrin Repeat Domain 30A* (*ANKRD30A*), *Alpha-Methylacyl-CoA Racemase* (*AMACR*), *Glucosaminyl (N-Acetyl) Transferase 1* (*GCNT1*), *PLA2G7*, *Androgen Receptor* (*AR*), *CPNE4*, *Periostin* (*POSTN*), *N-Acetylated Alpha-Linked Acidic Dipeptidase Like 2* (*NAALADL2*), and *ETS-Related Gene* (*ERG*), many of which are associated with immune evasion, hormonal signaling, and tumor progression. Conversely, significantly downregulated genes such as *Transmembrane Protease, Serine 5* (*TMPRSS5*), *Tripartite Motif Containing 9* (*TRIM9*), *Dolichyldiphosphatase 1* (*DOLPP1*), *Neuronal Guanine Nucleotide Exchange Factor* (*NGEF*), *Solute Carrier Family 13 Member 4* (*SLC13A4*), *Family With Sequence Similarity 124 Member A* (*FAM124A*), and *Leucine Rich Repeat Containing 61* (*LRRC61*) suggest a loss of regulatory functions within the tumor microenvironment. A detailed overview of the MF of all DEGs is presented in Table 3.

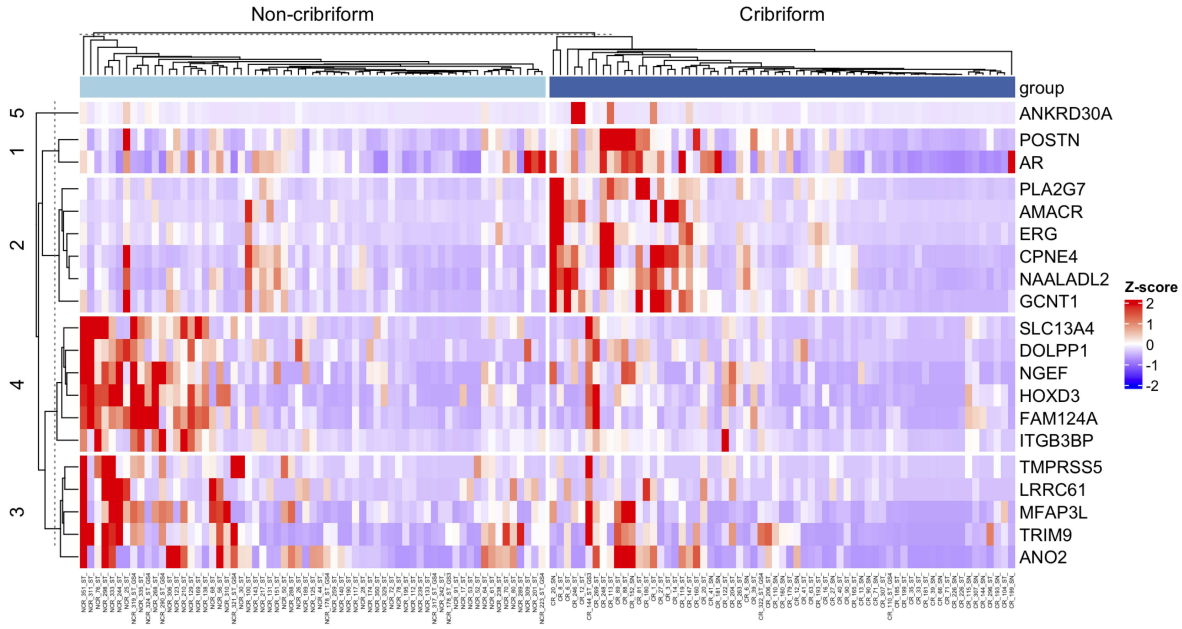
Table 3 – Top 20 DEGs in cribriform and non-cribriform PC.

Gene	Expression	Description	Reference
TMPRSS5	Underexpressed	Involved in viral entry and cell signaling.	[1]
TRIM9	Underexpressed	Plays a role in protein degradation and neuroinflammation.	[2]
DOLPP1	Underexpressed	Involved in glycosylation and cell surface protein processing.	[3]
NGEF	Underexpressed	Guanine nucleotide exchange factor, regulates Rho GTPases.	[4]
SLC13A4	Underexpressed	Sodium-dependent transporter involved in organic anion transport.	[5]
LRRC61	Underexpressed	A leucine-rich repeat protein, role in cellular signaling.	[6]
HOXD3	Underexpressed	Homeobox gene involved in development and differentiation.	[7]
ANO2	Underexpressed	Involved in calcium-activated chloride channels, cancer progression.	[8]
ITGB3BP	Underexpressed	Interacts with β 3 integrins; involved in adhesion and migration.	[18]
MFAP3L	Underexpressed	Involved in the regulation of the extracellular matrix.	[9]
FAM124A	Underexpressed	Associated with microtubule regulation.	[10]
GCNT1	Overexpressed	Glycosyltransferase involved in cancer metastasis and cell adhesion.	[11]
AMACR	Overexpressed	Involved in lipid metabolism, overexpressed in cancer.	[12]
PLA2G7	Overexpressed	Involved in phospholipid metabolism, inflammation.	[13]
AR	Overexpressed	Androgen receptor, critical in prostate cancer progression.	[14]
POSTN	Overexpressed	Involved in extracellular matrix remodeling and metastasis.	[15]
NAALADL2	Overexpressed	Involved in neuropeptide metabolism, linked to cancer progression.	[16]
ERG	Overexpressed	Transcription factor, fusion gene frequently found in prostate cancer.	[17]
CPNE4	Overexpressed	Involved in synaptic vesicle trafficking, potential cancer biomarker.	[19]
ANKRD30A	Overexpressed	Protein involved in cellular stress response and tumor progression.	[20]

Source: prepared by the author (2025).
Note: Gene expression was classified as overexpressed or underexpressed based on DEA results.

Additionally, a hierarchical clustering heatmap (Figure 22) highlights distinct gene expression patterns that clearly distinguish cribriform from non-cribriform samples.

Figure 22 – Heatmap illustrating the expression patterns of the top 20 DEGs) between cribriform and non-cribriform PC samples. Hierarchical clustering was applied to both genes and samples, highlighting distinct expression profiles between the two groups.



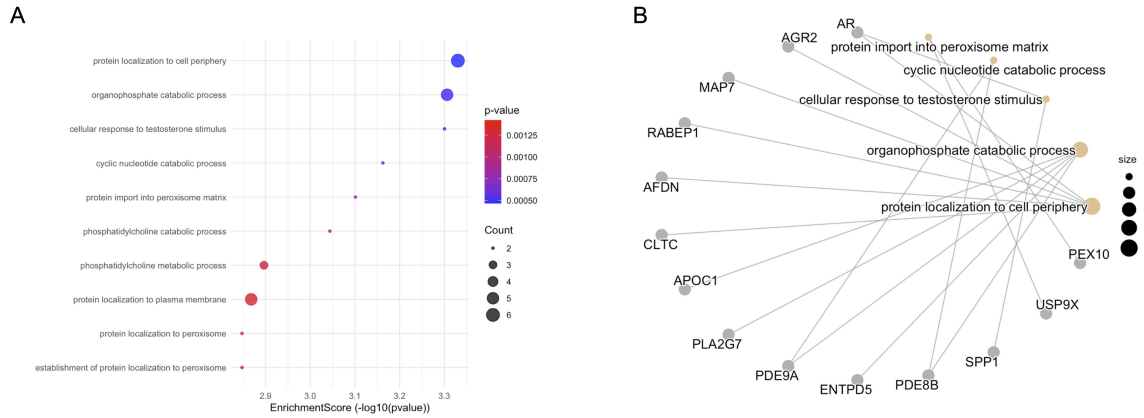
Source: prepared by the author (2025).

5.3.4 Gene Ontology and Pathway Analysis

The functional enrichment analysis (GO - BP) of DEGs) between the cribriform and non-cribriform patterns revealed distinct BP associated with each group. Upregulated processes in the CP included protein import into peroxisome matrix, cyclic nucleotide catabolic process, cellular response to testosterone stimulus, organophosphate catabolic process, and protein localization to cell periphery. These findings suggest a potential metabolic reprogramming, increased sensitivity to androgen signaling, and enhanced cellular adaptability, which may contribute to the aggressiveness of the cribriform phenotype.

The functional enrichment analysis (GO - MF) of upregulated genes in the CP revealed an enrichment in phosphodiesterase activities, including 3',5'-cyclic-AMP phosphodiesterase, 3',5'-cyclic-GMP phosphodiesterase, and cyclic-nucleotide phosphodiesterase activities (Figure 23). These findings suggest an increased degradation of cyclic nucleotides (cAMP and cGMP), which are known regulators of cell differentiation, proliferation, and apoptosis. The reduction of intracellular cAMP and cGMP levels has been associated with enhanced tumor aggressiveness and loss of differentiation, which aligns with the poorly differentiated nature of cribriform tumors.

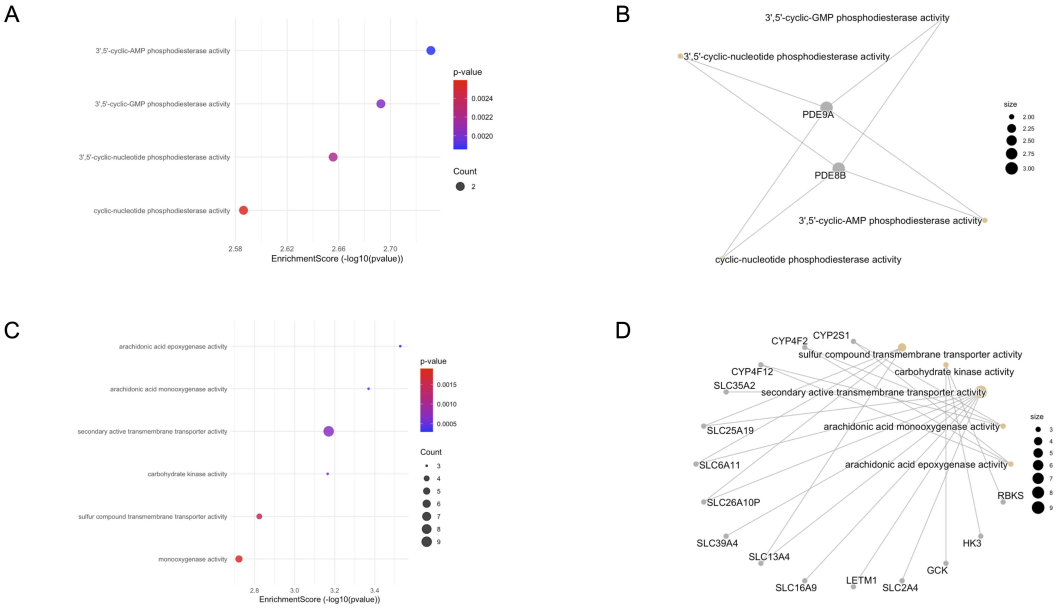
Figure 23 – GO enrichment analysis – BP for DEGs between cribriform and non-cribriform patterns. (A) Bubble plot showing enriched biological processes for upregulated genes. (B) Network representation of the interaction among upregulated genes within enriched biological processes.



Source: prepared by the author (2025).

In contrast, the downregulated functions in the CP included sulfur compound transmembrane transporter activity, carbohydrate kinase activity, and arachidonic acid monooxygenase/epoxygenase activities(Figure 24), suggesting potential alterations in metabolic pathways and oxidative stress regulation. These changes may contribute to the aggressive phenotype of the CP by promoting a metabolic shift that supports tumor progression and survival.

Figure 24 – Functional enrichment analysis (GO - MF) of DEGs. (A) Bubble plot representing the enriched MF of upregulated genes in the CP. (B) Network representation of the upregulated gene interactions within enriched MF. (C) Bubble plot depicting the enriched molecular functions of downregulated genes. (D) Network representation of the downregulated gene interactions within enriched molecular functions.



Source: prepared by the author (2025).

5.4 Discussion

CP has been consistently associated with adverse clinical outcomes in PC, including higher risks of recurrence, metastasis, and cancer-specific mortality (Oufattole *et al.*, 2023). In light of these associations, recent studies have aimed to elucidate the underlying molecular mechanisms and to identify biomarkers that may provide insight into the link between this histological pattern and the aggressive behavior observed in prostatic adenocarcinomas.

Previous studies, such as Wong *et al.* (2022) and Chappidi *et al.* (2024), employed distinct approaches to explore the biology of the CP. Wong *et al.* used single-cell RNA-seq on paired tumor and benign samples, identifying pathways linked to metastasis and an immuno-suppressive microenvironment in cribriform lesions. Chappidi *et al.* analyzed transcriptomic heterogeneity among GP4 subtypes (EC and IDC) using GRID data and clustering methods, revealing associations with steroid hormone signaling and biochemical recurrence (Hazard Ratio (HR) = 2.35; 95% CI = 1.01–5.47). In contrast, our study applied DEA using Next-Generation Sequencing (NGS) on a larger cohort of 130 samples evenly distributed between cribriform and non-cribriform cases, enhancing statistical power and enabling a more accurate tissue-level

transcriptomic characterization to improve risk stratification.

To date, no studies in the literature have employed DEA using NGS data in a robust cohort of 130 PC samples, specifically focusing on the association between CP and poorer disease prognosis. These findings underscore the novelty of the present study and highlight the importance of investigating molecular pathways that remain inadequately elucidated in this context.

In this study, we performed a gene expression profiling analysis in PC samples with and without CP, using publicly available transcriptomic data from the GEO portal NCBI. As a reference, we used the dataset from the study conducted by Ma *et al.* (2023), which analyzed 293 tumor and benign stromal samples using RNA-seq. In that study, DEA was performed separately for different profiles: low ($\leq$ 6 and 3 + 4) and high (4 + 3 and 8-10 Gleason scores, presence of CP (in samples with Gleason 7 to 10), and between lethal and indolent cases. However, the study presented an unbalanced distribution between groups, with only 45 samples showing CP and 195 without it, which may have compromised the robustness of the analysis.

Several studies have emphasized the impact of sample imbalance on DEA. Yang *et al.* (2006) demonstrated that discrepancies in sample size between comparison groups can compromise the identification of DEGs, affecting the statistical performance of the methods used. Complementarily, Sprang *et al.* (2024) showed that not only the quantity but also the quality of the samples can influence the validity of the results. By analyzing clinical RNA-seq datasets, the authors observed that datasets with disproportionate sample quality between groups yielded a higher number of DEGs with low biological relevance and reduced reproducibility.

Simply removing samples from the control group to match the number of cases with CP would not be appropriate, as this approach could introduce selection bias and compromise the integrity of subsequent analyses (Maan *et al.*, 2022). To address this limitation, we applied a clustering strategy to the control group (non-cribriform), based on the Euclidean distance of their gene expression profiles, followed by stratified sampling. This method preserved the intrinsic biological variability within the control group and yielded a balanced dataset composed of 45 cribriform and 45 non-cribriform samples. Although stratified sampling based on Euclidean distance is not yet widely reported in the literature, existing studies emphasize the importance of accounting for inter-sample distances to reduce bias and enhance result validity, particularly in imbalanced datasets. Compared to traditional numerical downsampling, this approach provided a more representative sample selection and enabled more robust and reliable statistical analysis.

To overcome this limitation, we applied a clustering strategy to the control group (non-cribriform), based on the Euclidean distance between their gene expression profiles, followed by stratified sampling (Ultsch *et al.*, 2022). This approach ensured that the selected samples preserved the intrinsic biological variability of the control group, resulting in a balanced dataset of 45 cribriform and 45 non-cribriform samples. Although not widely documented, stratified sampling guided by Euclidean distance has been suggested as a means to reduce bias and improve the validity of results in imbalanced datasets (Tsai *et al.*, 2016). Compared to traditional numeric equalization methods, this strategy provided a more representative sample distribution and enabled more robust and reliable statistical analysis (Ultsch *et al.*, 2022).

Our analysis identified 55 DEGs ($p < 0.01$) between cribriform and non-cribriform PC samples, compared to 56 genes in the original study. Notably, the previous analysis used the limma package (v3.48.3) with voom transformation, which assumes normality and variance homogeneity—conditions not met given their unbalanced groups (45 vs. 195 samples) (Ritchie *et al.*, 2015; Law *et al.*, 2014). Without applying balancing strategies, such an imbalance may lead to unstable variance estimates, reducing the statistical power and reliability of DEGs detection (You *et al.*, 2023). To address limma's limitations, we used edgeR, which models raw count data using a negative binomial distribution (Robinson *et al.*, 2010). Its robustness in small, balanced sample sets—achieved here via stratified sampling after clustering—enhanced statistical precision, improving the detection of DEGs, including low-abundance transcripts (Varet *et al.*, 2016; Petrany *et al.*, 2024).

From the Venn diagram analysis, we identified three DEGs in both studies: *CPNE4*, *PTPRJ*, and *PLA2G7*. *CPNE4*, a calcium-dependent membrane-associated protein involved in neural development, has been reported as overexpressed in neuroendocrine and small cell prostate cancers, suggesting a role in prostate tumor progression (Tsai *et al.*, 2017). *PTPRJ* (DEP-1) encodes a tyrosine phosphatase with an ambiguous role in various tumor cancer, acting either as a tumor suppressor or oncogene (Sivaganesh *et al.*, 2021). *PLA2G7* is a phospholipase associated with platelet-activating factor and oxidized phospholipids, serving as a prognostic biomarker in PC, lymphoma, and melanoma (Xie *et al.*, 2024; Vainio *et al.*, 2011). Its silencing in diffuse large B-cell lymphoma cells reduces proliferation and migration, and induces apoptosis (Zheng *et al.*, 2021).

In the DEA between cribriform and non-cribriform PC samples, we identified a set of genes significantly underexpressed in the cribriform group, potentially playing a role in sup-

pressing aggressive disease features. Notable genes include *TMPRSS5, TRIM9, DOLPP1, NGEF*, and *SLC13A4*, involved in regulating metabolic pathways, extracellular matrix remodeling, and cell signaling—processes often altered during tumor progression (Zhang *et al.*, 2023; Mi *et al.*, 2025; Yang *et al.*, 2021; Bugge *et al.*, 2009; Mamoor *et al.*, 2023). Additionally, genes such as *HOXD3* and *ANO2*, linked to cell differentiation and ionic homeostasis, respectively, show underexpression that may reflect a loss of homeostatic control in aggressive morphologies (Yang *et al.*, 2019; Tengvall *et al.*, 2019). Reduced expression of *ITGB3BP* and *MFAP3L* suggests a role in modulating cell adhesion and matrix integrity, often compromised in cribriform tumors (Liu *et al.*, 2022; Ye *et al.*, 2024). These findings indicate that the CP may be associated with a distinct transcriptional profile, characterized by the repression of genes involved in regulating the tumor microenvironment and containing the aggressive phenotype.

Among the genes significantly overexpressed in samples with CP, *GCNT1*, *AMACR*, *PLA2G7*, *AR*, and *ERG* stand out, all previously reported in PC progression. *GCNT1* is involved in the glycosylation of mucins, associated with immune evasion and metastasis (Hodgson *et al.*, 2023); *AMACR*, a known marker for prostate carcinoma, plays a role in fatty acid metabolism (Alinezhad *et al.*, 2016); and *PLA2G7* is linked to inflammation and tissue remodeling (Vainio *et al.*, 2011). The overexpression of AR and ERG highlights the role of hormonal and transcriptional pathways in the CP (Zengin *et al.*, 2024; Khosh Kish *et al.*, 2022). Genes such as *POSTN* and *NAALADL2* have been associated with tumor invasiveness (Cattrini *et al.*, 2020; Groen *et al.*, 2022), while *CPNE4* and *ANKRD30A* may be related to alterations in cell adhesion and extracellular matrix organization (Pariyar *et al.*, 2022; Tang *et al.*, 2021), characteristics typical of the cribriform phenotype. These findings suggest that the CP is associated with a pro-tumorigenic transcriptional reprogramming and a more aggressive tumor microenvironment.

The BP enrichment analysis of upregulated genes in cribriform PC samples revealed an association with tumor aggressiveness-related processes. The "protein localization to cell periphery" process suggests that these proteins may be involved in forming and maintaining the invasive characteristics of tumor cells, a hallmark of cribriform cancer aggressiveness (Lee *et al.*, 2018; Destouni *et al.*, 2022). Additionally, processes such as "organophosphate catabolic process" and "cellular response to testosterone stimulus" point to potential mechanisms of hormonal resistance, commonly observed in aggressive cancers (Blatt *et al.*, 2021). "Cyclic nucleotide catabolic process" and "protein import into peroxisome matrix" suggest intensified cellular modulation aimed at maintaining homeostasis in a tumorigenic microenvironment (Gao

et al., 2022; Fajardo *et al.*, 2014). These findings suggest a reprogrammed transcriptional profile linked to tumor progression in cribriform PC.

The functional enrichment analysis of upregulated genes in cribriform PC samples revealed a significant concentration of molecular functions associated with cyclic phosphodiesterase activity, such as "3',5'-cyclic-AMP phosphodiesterase activity" and "cyclic-nucleotide phosphodiesterase activity", which degrade cyclic nucleotides like cAMP and cGMP, regulating cellular signaling pathways (Montoya-Durango *et al.*, 2023). This upregulation suggests a negative regulation of processes like growth, apoptosis, and differentiation, contributing to the aggressiveness and invasiveness of cribriform cancer, as well as being linked to tumor progression and hormonal resistance (Hesterberg *et al.*, 2021). In contrast, the analysis of downregulated genes indicated the loss of functions such as "arachidonic acid epoxygenase activity" and "monooxygenase activity", involved in cellular signaling and metabolite transport (Apaya *et al.*, 2019). The downregulation of these enzymes may impair eicosanoid production and compromise cellular homeostasis, fostering a more deregulated and aggressive tumor microenvironment, which are characteristic features of cribriform prostate carcinoma (Ershov *et al.*, 2025; Masamrekh *et al.*, 2020).

5.5 Conclusion

The present work advances our understanding of the transcriptomic landscape associated with CP in PC. Using a balanced dataset of 90 cases and robust methodological strategies—including clustering by Euclidean distance followed by stratified sampling. DEGs not only confirmed previously reported biomarkers (e.g., *PLA2G7, CPNE4, PTPRJ*) but also revealed novel candidates potentially linked to tumor progression, immune evasion, and remodeling of the tumor microenvironment. The upregulation of genes such as *GCNT1*, *AMACR*, *AR*, and *ERG* underscores the involvement of hormonal and inflammatory pathways in the aggressive behavior of cribriform tumors, while downregulation of *TMPRSS5*, *TRIM9*, and *HOXD3* points to a loss of homeostatic and tumor-suppressive functions. Functional enrichment analyses revealed disruptions in processes and MF related to hormonal resistance, cytoskeletal reorganization, lipid metabolism, and cell adhesion, reinforcing the classification of CP as a distinct molecular subtype of PC with more aggressive clinical features. These findings lay the groundwork for the identification of prognostic biomarkers and therapeutic targets tailored to this high-risk subgroup.

6 FINAL CONSIDERATIONS

The present thesis investigated the morphological and molecular determinants of tumor aggressiveness in PC, with a special focus on cribriform morphology. By integrating digital pathology techniques and transcriptomic analysis, we provided a multidimensional perspective that reinforced the role of this histological pattern as a marker of poor prognosis.

In the first article, the use of digital pathology tools allowed for objective quantification of histological and immunohistochemical features across 72 prostatectomy samples. The study highlighted a consistent association between cribriform morphology and aggressive clinical-pathological features such as higher pathological stage and lack of organ confinement. A logistic regression model combining GP4 percentage and ATM expression yielded robust predictive power for cribriform morphology (AUC = 0.79), demonstrating the potential of integrated histological and molecular analysis for risk stratification. In the second article, an in silico transcriptomic analysis of 90 samples—balanced between cribriform and non-cribriform tumors—identified 55 DEGs. This transcriptional signature included overexpressed genes involved in immune evasion, hormonal signaling, and lipid metabolism, as well as underexpressed genes related to cellular adhesion and extracellular matrix regulation. Functional enrichment analyses further supported a transcriptional reprogramming toward a more aggressive phenotype in cribriform tumors, contributing to a better understanding of the molecular basis of their unfavorable behavior.

Together, the two studies converge to characterize cribriform morphology not merely as a structural variant, but as a phenotype that integrates histological, molecular, and transcriptional features of tumor aggressiveness. The findings contribute to the growing body of evidence advocating for the routine recognition of cribriform architecture in diagnostic practice and support the development of molecular tools to refine prognostic stratification.

Although based on different sample sets and methodological designs, the two chapters of this thesis are conceptually complementary and converge toward a shared objective: the elucidation of the biological foundations of cribriform morphology in PC. The use of independent cohorts strengthens the generalizability of the findings and provides a broader view of this architectural pattern across different analytical contexts. This integrated approach reinforces the importance of combining morphological evaluation with molecular data to uncover consistent biological signatures, validate potential biomarkers, and support more accurate prognostic models.

Future research should expand the sample size and validate the proposed biomarkers in independent cohorts. Additionally, the incorporation of spatial transcriptomics, multiplexed imaging, and artificial intelligence could further improve the precision and clinical utility of morpho-molecular evaluations. Overall, this thesis reinforces the value of integrated pathology and genomics in unveiling the biological complexity of PC and advancing personalized oncological care.

REFERENCES

ALINEZHAD, Saeid *et al.* Global expression of AMACR transcripts predicts risk for prostate cancer – a systematic comparison of AMACR protein and mRNA expression in cancerous and noncancerous prostate. **BMC Urology**, v. 16, n. 1, p. 10, 2016. Available from: https://doi.org/10.1186/s12894-016-0128-8. Accessed on: 14 jul. 2025.

ALBUQUERQUE, C. G. P. *et al.* PTEN loss in Gleason grade 7 prostate tumors exhibits intratumoral heterogeneity and is associated with unfavorable pathological features. **Applied Cancer Research**, v. 39, n. 1, p. 1, 2019. Available from: https://doi.org/10.1186/s41241-018-0071-y. Accessed on: 03 apr. 2025.

AMERICAN CANCER SOCIETY. Prostate cancer is number 1 for 118 countries worldwide. 2025. Available from: https://www.cancer.org/research/acs-research-news/prostate-cancer-is-number-1-for-118-countries-worldwide.html. Accessed on: 2 aug. 2025.

APAYA, Maria Karmella *et al.* Integrated omics-based pathway analyses uncover CYP epoxygenase-associated networks as theranostic targets for metastatic triple negative breast cancer. **Journal of Experimental & Clinical Cancer Research**, v. 38, n. 1, p. 187, dez. 2019. Available from: https://doi.org/10.1186/s13046-019-1187-y. Accessed on: 19 jul. 2025.

ASIF, A. *et al.* Unleashing the potential of AI for pathology: challenges and considerations in clinical deployment. **The Journal of Pathology**, 2023. Disponível em: https://doi.org/10.1002/path.6168. Acesso em: 10 jul. 2025.

BANKHEAD, P. *et al.* QuPath: open source software for digital pathology image analysis. **Scientific Reports**, v. 7, n. 1, p. 16878, 2017. Available from: https://doi.org/10.1038/s41598-017-17204-5. Accessed on: 08 apr. 2025.

BARKER, C. *et al.* Digital pathology: integration of whole-slide imaging in routine diagnostic workflows. **Journal of Pathology Informatics**, v. 15, p. 100408, 2024. Available from: https://doi.org/10.1016/j.jpi.2024.100408. Accessed on: 2 aug. 2025.

BARBIERI, C. E. *et al.* The mutational landscape of prostate cancer. **European Urology**, v. 68, n. 5, p. 795–804, 2021. Available from: https://doi.org/10.1016/j.eururo.2020.12.003. Accessed on: 10 jul. 2025.

BASAK, K. *et al.* Whole slide images in artificial intelligence applications in digital pathology: challenges and pitfalls. **Frontiers in Artificial Intelligence**, v.6, 2023. Disponível em: https://doi.org/10.3389/frai.2023.xxxxxx. Acesso em: 10 jul. 2025.

BHOLOWALIA, Purnima; KUMAR, Arvind. EBK-Means: A clustering technique based on elbow method and K-means in WSN. **International Journal of Computer Applications**, v. 105, n. 9, p. 17–24, nov. 2014. Available from: https://doi.org/10.5120/18405-9674. Accessed on: 19 jul. 2025.

BLATT, Eliot B. *et al.* Overcoming oncogene addiction in breast and prostate cancers: a comparative mechanistic overview. **Endocrine-Related Cancer**, v. 28, n. 2, p. R31-R46, fev. 2021. Available from: https://doi.org/10.1530/ERC-20-0272. Accessed on: 19 jul. 2025.

BOGAARD, Mari. *et al.* High proliferative cribriform prostate cancer defines a patient subgroup with an inferior prognosis. **Histopathology**, v. 83, n. 6, p. 853–869, 2023. Available from: https://doi.org/10.1111/his.15012. Accessed on: 02 apr. 2025.

BOGAARD, Mari *et al.* GRIN3A: a biomarker associated with a cribriform pattern and poor prognosis in prostate cancer. **Neoplasia (United States)**, v. 55, 2024. Available from: https://doi.org/10.1016/j.neo.2024.101023. Accessed on: 12 jul. 2025.

BOTTCHER, René *et al.* Cribriform and intraductal prostate cancer are associated with increased genomic instability and distinct genomic alterations. **BMC Cancer**, v. 18, n. 1, 2018. Available from: https://doi.org/10.1186/s12885-017-3976-z. Accessed on: 12 jul. 2025.

BREA, L. *et al.* Tumor-intrinsic regulators of the immune-cold microenvironment of prostate cancer. **Trends in Endocrinology & Metabolism**, 2025. Available from: https://doi.org/10.1016/j.tem.2024.12.003. Accessed on: 09 apr. 2025.

BUGGE, Thomas H.; ANTALIS, Toni M.; WU, Qingyu. Type II Transmembrane Serine Proteases. **Journal of Biological Chemistry**, v. 284, n. 35, p. 23177–23181, ago. 2009. Available from: https://doi.org/10.1074/jbc.R109.021006. Accessed on: 19 jul. 2025.

BULTEN, W. *et al.* Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. **Modern Pathology**, v. 34, p. 660–671, 2021. Disponível em: https://doi.org/10.1038/s41379-020-0640-y. Acesso em: 10 jul. 2025.

CATTRINI, Carlo *et al.* Integrative Analysis of Periostin in Primary and Advanced Prostate Cancer. **Translational Oncology**, v. 13, n. 7, p. 100789, jul. 2020. Available from: https://doi.org/10.1016/j.tranon.2020.100789. Accessed on: 14 jul. 2025.

CHAPPIDI, Meera R. *et al.* Transcriptomic heterogeneity of expansile cribriform and other Gleason pattern 4 prostate cancer subtypes. **European Urology Oncology**, v. 7, n. 2, p. 222–230, abr. 2024. Available from: https://doi.org/10.1016/j.euo.2023.06.007. Accessed on: 14 jul. 2025.

CHEN, Z. *et al.* Prognostic value of cribriform size, percentage, and intraductal carcinoma in Gleason score 7 prostate cancer with cribriform Gleason pattern 4. **Human Pathology**, v. 118, p. 18–29, 2021. Available from: https://doi.org/10.1016/j.humpath.2021.09.005. Accessed on: 08 apr. 2025.

CORCHETE, L. A. *et al.* Systematic comparison and assessment of RNA-seq protocols reveals high concordance with qRT-PCR as a gold standard. **Scientific Reports**, v. 10, n. 1, 2020. Disponível em: https://doi.org/10.1038/s41598-020-76881-x. Acesso em: 10 jul. 2025.

COSTA-SILVA, J. *et al.* Computational methods for differentially expressed gene analysis from RNA-Seq: an overview. **PeerJ**, v. 9, eXXXX, 2021. Disponível em: https://doi.org/10.1101/2021.09.08.XXXX. Acesso em: 10 jul. 2025.

CUI, Z. *et al.* Prostate cancer: epidemiology, risk factors, and molecular mechanisms. **Cancers**, v. 16, n. 2, p. 287, 2024. Available from: https://doi.org/10.3390/cancers16020287. Accessed on: 10 jul. 2025.

DALLANG, B. *et al.* Prognostic significance of Ki67 expression in prostate cancer in Nigerians: a single-center study. **Cureus**, 2025. Available from: https://doi.org/10.7759/cureus.80997.

Accessed on: 27 mar. 2025.

DE-CHAO, F. *et al.* Current best practices in single-cell RNA-Seq analysis for prostate cancer: from sequencing to bioinformatic integration. **Molecular Systems Biology Review**, v. 16, n. 1, 2024. Available from: https://doi.org/10.1186/s40779-024-00526-7. Accessed on: 10 jul. 2025.

DESTOUNI, Maria; LAZARIS, Andreas C.; TZELEPI, Vasiliki. Cribriform Patterned Lesions in the Prostate Gland with Emphasis on Differential Diagnosis and Clinical Significance. **Cancers**, v. 14, n. 13, p. 3041, jun. 2022. Available from: https://doi.org/10.3390/cancers14133041. Accessed on: 12 jul. 2025.

DONG, Fei *et al.* Architectural heterogeneity and cribriform pattern predict adverse clinical outcome for Gleason grade 4 prostatic adenocarcinoma. **The American Journal of Surgical Pathology**, v. 37, n. 12, p. 1855–1861, 2013.

ELFANDY, Habiba *et al.* Genetic and epigenetic determinants of aggressiveness in cribriform carcinoma of the prostate. **Molecular Cancer Research**, v. 17, n. 2, p. 446–456, 2019. Available from: https://doi.org/10.1158/1541-7786.MCR-18-0440. Accessed on: 12 jul. 2025.

EPSTEIN, J. I. *et al.* A contemporary prostate cancer grading system: a validated alternative to the Gleason score. **European Urology**, v. 69, n. 3, p. 428–435, mar. 2016. Disponível em: https://doi.org/10.1016/j.eururo.2015.06.046. Acesso em: 14 jul. 2025.

ERSHOV, Pavel V. *et al.* Protein subinteractomes of human microsomal cytochromes P450. **Molecular Biology Reports**, v. 52, n. 1, p. 226, dez. 2025. Available from: https://doi.org/10.1007/s11033-025-10341-5. Accessed on: 12 jul. 2025.

FAJARDO, Alexandra; PIAZZA, Gary; TINSLEY, Heather. The Role of Cyclic Nucleotide Signaling Pathways in Cancer: Targets for Prevention and Treatment. **Cancers**, v. 6, n. 1, p. 436-458, fev. 2014. Available from: https://doi.org/10.3390/cancers6010436. Accessed on: 12 jul. 2025.

FALAHATKAR, *et al.*. An update on biochemical and genomic markers for prostate cancer. **Urology journal** (2021) Available from: https://doi.org/10.22037/uj.v18i.6828. Accessed on: 06 apr. 2025.

FENG, D.-C. *et al.* The implications of single-cell RNA-Seq analysis in prostate cancer: unraveling tumor heterogeneity and therapeutic pathways. **Military Medical Research**, v. 11, n. 1, 2024. Available from: https://doi.org/10.1186/s40779-024-00526-7. Accessed on: 10 jul. 2025.

GAO, Yuan *et al.* Protein import into peroxisomes occurs through a nuclear pore–like phase. **Science**, v. 378, n. 6625, dez. 2022. Available from: https://doi.org/10.1126/science.adf3971. Accessed on: 12 jul. 2025.

GENEVIA TECHNOLOGIES. RNA-Seq data analysis. [S.l.], s.d. Available from: https://geneviatechnologies.com/bioinformatics-analyses/rna-seq-data-analysis/. Accessed on: 2 aug. 2025.

GREENE, Christine *et al.* The Potential of Digital Image Analysis to Determine Tumor Cell Content in Biobanked Formalin-Fixed, Paraffin-Embedded Tissue Samples. **Biopreservation and**

Biobanking, v. 19, n. 4, p. 324–331, 2021. Available from: https://doi.org/10.1089/bio.2020.0105. Accessed on: 05 jul. 2025.

GROEN, Levi *et al.* The Androgen Regulated lncRNA NAALADL2-AS2 Promotes Tumor Cell Survival in Prostate Cancer. **Non-Coding RNA**, v. 8, n. 6, p. 81, dez. 2022. Available from: https://doi.org/10.3390/ncrna8060081. Accessed on: 12 jul. 2025.

HAFFNER, M. C. *et al.* Tracking the clonal origin of lethal prostate cancer. **Journal of Clinical Investigation**, v. 131, n. 12, p. e149231, 2021. Available from: https://doi.org/10.1172/JCI149231. Accessed on: 10 jul. 2025.

HAM, Won Sik *et al.* The impact of downgrading from biopsy Gleason 7 to prostatectomy Gleason 6 on biochemical recurrence and prostate cancer specific mortality. **Journal of Urology**, v. 197, n. 4, p. 1060-1067, 2017. Available from: https://doi.org/10.1016/j.juro.2016.11.079. Accessed on: 12 jul. 2025.

HE, M. *et al.* Research progress on deep learning in magnetic resonance imaging–based diagnosis and treatment of prostate cancer: a review on the current status and perspectives. **Frontiers in Oncology**, v. 13, 2023. Available from: https://doi.org/10.3389/fonc.2023.1189370. Accessed on: 12 jul. 2025.

HELPAP, B. *et al.* The significance of accurate determination of Gleason score for therapeutic options and prognosis of prostate cancer. **PathologyOncology Research**, v. 22, n. 2, p. 349–356, abr. 2016. Disponível em: https://doi.org/10.1007/s12253-015-0013-x. Acesso em: 10 jul. 2025.

HEIDELBAUGH, J. J.; HADJ-MOUSSA, M. Men's health: prostate cancer screening. **FP Essentials**, v. 503, p. 11–17, 2021. Available from: https://pubmed.ncbi.nlm.nih.gov/33856178/. Accessed on: 12 jul. 2025.

HESTERBERG, Amanda B.; GORDETSKY, Jennifer B.; HURLEY, Paula J. Cribriform Prostate Cancer: Clinical Pathologic and Molecular Considerations. **Urology**, v. 155, p. 47–54, set. 2021. Available from: https://doi.org/10.1016/j.urology.2021.05.028. Accessed on: 12 jul. 2025.

HIJAZI, A. *et al.* Digital pathology for better clinical practice. **Cancers**, v. 16, n. 9, p. 1686, 2024. Available from: https://doi.org/10.3390/cancers16091686. Accessed on: 12 jul. 2025.

HODGSON, Kirsty *et al.* The role of GCNT1 mediated O-glycosylation in aggressive prostate cancer. **Scientific Reports**, v. 13, n. 1, p. 17031, out. 2023. Available from: https://doi.org/10.1038/s41598-023-43019-8. Accessed on: 12 jul. 2025.

HOLLEMANS, Eva *et al.* Large cribriform growth pattern identifies ISUP grade 2 prostate cancer at high risk for recurrence and metastasis. **Modern Pathology**, v. 32, n. 1, p. 139-146, 2019. Available from: https://doi.org/10.1038/s41379-018-0157-9. Accessed on: 12 jul. 2025.

HOMMERDING, O. *et al.* High interobserver variability of PTEN immunohistochemistry defining PTEN status in low- to intermediate-risk prostate cancer: results of the first German ring trial. **Virchows Archiv**, 2024. Available from: https://doi.org/10.1007/s00428-024-03999-y. Accessed on: 15 may 2025.

INSTITUTO NACIONAL DE CÂNCER (INCA). **Estimativa** 2023: incidência de câncer no Brasil. Rio de Janeiro: INCA, 2022. Available from: https://www.inca.gov.br/estimativa.

Accessed on: 10 jul. 2025.

JAMASPISHVILI, T. *et al.* Clinical implications of PTEN loss in prostate cancer. **Nature Reviews Urology**, v. 15, n. 4, p. 222–234, 2018. Available from: https://doi.org/10.1038/nrurol.2018.9. Accessed on: 04 apr. 2025.

KARWACKI, J. *et al.* Association of lymphovascular invasion with biochemical recurrence and adverse pathological characteristics of prostate cancer: a systematic review and meta-analysis. **European Urology Open Science**, v. 69, p. 112–126, 2024. Available from: https://doi.org/10.1016/j.euros.2024.09.007. Accessed on: 02 abr. 2025.

KAUR, H. *et al.* Genomic and clinicopathologic characterization of ATM-deficient prostate cancer. **Clinical Cancer Research**, v. 26, n. 18, p. 4869–4881, 2020. Available from: https://doi.org/10.1158/1078-0432.CCR-20-0764. Accessed on: 12 apr. 2025.

KHOSH KISH, Ealia *et al.* The Expression of Proto-Oncogene ETS-Related Gene (ERG) Plays a Central Role in the Oncogenic Mechanism Involved in the Development and Progression of Prostate Cancer. **International Journal of Molecular Sciences**, v. 23, n. 9, p. 4772, abr. 2022. Available from: https://doi.org/10.3390/ijms23094772. Accessed on: 12 jul. 2025.

KLEIN, R. J. *et al.* Prostate cancer polygenic risk score and prediction of lethal prostate cancer. **npj Precision Oncology**, v. 6, n. 1, p. 25, 2022. Available from: https://doi.org/10.1038/s41698-022-00266-8. Accessed on: 12 jul. 2025.

KROON, L. J. *et al.* Comparison of highest and overall percentage Gleason pattern 4 in prostate cancer biopsies. **Virchows Archiv**, 2025. Available from: https://doi.org/10.1007/s00428-025-04117-2. Accessed on: 03 apr. 2025.

KULESHOV, Maxim V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. **Nucleic Acids Research**, v. 44, n. W1, p. W90–W97, 2016. Available from: https://doi.org/10.1093/nar/gkw377. Accessed on: 12 jul. 2025.

KWELDAM, Charlotte F. *et al.* Cribriform growth is highly predictive for postoperative metastasis and disease-specific death in Gleason score 7 prostate cancer. **Modern Pathology**, v. 28, n. 3, p. 457-464, 2015. Available from: https://doi.org/10.1038/modpathol.2014.116. Accessed on: 12 jul. 2025.

KWELDAM, C. F. *et al.* On cribriform prostate cancer. **Translational Andrology and Urology**, v. 7, n. 1, p. 145–154, 2018. Disponível em: https://doi.org/10.21037/tau.2017.12.33. Acesso em: 10 jul. 2025.

KWON, W. A. *et al.* Immunotherapy in prostate cancer: from a “cold” tumor to a “hot” prospect. **Cancers**, v. 17, n. 7, p. 1064, 2025. Available from: https://doi.org/10.3390/cancers17071064. Accessed on: 12 apr. 2025.

LAW, Charity W.; CHEN, Yunshun; SHI, Wei; SMYTH, Gordon K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. **Genome Biology**, v. 15, n. 2, p. R29, 2014. Available from: https://doi.org/10.1186/gb-2014-15-2-r29. Accessed on: 12 jul. 2025.

LEE, S. H. *et al.* Universal immunohistochemistry analyzer for standardized and quantitative IHC assessment using deep learning. **Cell Reports Medicine**, v. 5, n. 4, 2024. Available from:

https://doi.org/10.1038/s41698-024-00770-z. Accessed on: 10 jul. 2025.

LEE, Thomas K.; RO, Jae Y. Spectrum of Cribriform Proliferations of the Prostate: From Benign to Malignant. **Archives of Pathology & Laboratory Medicine**, v. 142, n. 8, p. 938-946, ago. 2018. Available from: https://doi.org/10.5858/arpa.2018-0005-RA. Accessed on: 12 jul. 2025.

LEITÃO, C. *et al.* Perceptions, knowledge, and attitudes of general population about prostate cancer-associated risk factors: a systematic review of qualitative studies focusing on lifestyle. **Current Oncology Reports**, v. 27, n. 4, p. 375–389, 2025. Available from: https://doi.org/10.1007/s11912-025-01653-7. Accessed on: 26 mar. 2025.

LIU, Zhendong *et al.* ITGB3BP is a potential biomarker associated with poor prognosis of glioma. **Journal of Cellular and Molecular Medicine**, v. 26, n. 3, p. 813–827, fev. 2022. Available from: https://doi.org/10.1111/jcmm.17127. Accessed on: 12 jul. 2025.

LONE, Zaeem *et al.* Transcriptomic features of cribriform and intraductal carcinoma of the prostate. **European Urology Focus**, v. 8, n. 6, p. 1575–1582, nov. 2022. Available from: https://doi.org/10.1016/j.euf.2022.05.005. Accessed on: 12 jul. 2025.

LOPES, A.; WARD, A. D.; CECCHINI, M. Eye tracking in digital pathology: a comprehensive literature review. **Journal of Pathology Informatics**, v. 15, p. 100383, 2024. Available from: https://doi.org/10.1016/j.jpi.2024.100383. Accessed on: 12 jul. 2025.

LOTAN, T. L. *et al.* Morphologic patterns and molecular features of prostate cancer. 2017. Available from: https://doi.org/10.48550/arXiv.1705.02678. Accessed on: 2 aug. 2025.

LOZANO, Rebeca *et al.* Association between BRCA2 alterations and intraductal and cribriform histologies in prostate cancer. **European Journal of Cancer**, v. 147, p. 74–83, abr. 2021. Available from: https://doi.org/10.1016/j.ejca.2021.01.027. Accessed on: 12 jul. 2025.

MA, C. *et al.* The prostate stromal transcriptome in aggressive and lethal prostate cancer. **Molecular Cancer Research**, 2023. Available from: https://doi.org/10.1158/1541-7786.MCR-23-XXXX. Accessed on: 10 jul. 2025.

MAAN, Hassaan *et al.* The differential impacts of dataset imbalance in single-cell data integration. [Pré-print]. 2022. Available from: https://doi.org/10.1101/2022.10.06.511156. Accessed on: 20 jul. 2025.

MAMOOR, Shahan. DOLPP1 is differentially expressed in central nervous system metastasis in human breast cancer. **Preprint [osf.io]**, 2023. Available from: https://doi.org/10.31219/osf.io/2gwdc. Accessed on: 20 jul. 2025.

MASAMREKH, Rami *et al.* Estimation of the inhibiting impact of abiraterone D4A metabolite on human steroid 21-monooxygenase (CYP21A2). **Steroids**, v. 154, p. 108528, fev. 2020. Available from: https://doi.org/10.1016/j.steroids.2019.108528. Accessed on: 20 jul. 2025.

MI, Jie *et al.* Neuronal guanine nucleotide exchange factor promotes the axonal growth and cancer cell proliferation via Ephrin-A3/EphA2 axis in lung adenocarcinoma. **Journal of Translational Medicine**, v. 23, n. 1, p. 246, fev. 2025. Available from: https://doi.org/10.1186/s12967-025-06233-8. Accessed on: 20 jul. 2025.

MIURA, Noriyoshi *et al.* The prognostic impact of intraductal carcinoma of the prostate: a systematic review and meta-analysis. **Journal of Urology**, v. 204, n. 5, p. 909-917, 2020. Available from: https://doi.org/10.1097/JU.0000000000001290. Accessed on: 12 jul. 2025.

MONTOYA-DURANGO, Diego *et al.* Dysregulated Cyclic Nucleotide Metabolism in Alcohol-Associated Steatohepatitis: Implications for Novel Targeted Therapies. **Biology**, v. 12, n. 10, p. 1321, out. 2023. Available from: https://doi.org/10.3390/biology12101321. Accessed on: 25 jul. 2025.

MORATIN, J. *et al.* Digital pathology scoring of immunohistochemical staining reliably identifies prognostic markers and anatomical associations in a large cohort of oral cancers. **Frontiers in Oncology**, v. 11, 2021. Available from: https://doi.org/10.3389/fonc.2021.712944. Accessed on: 05 may 2025.

MUN, Y. *et al.* Yet Another Automated Gleason Grading System (YAAGGS): using deep learning to predict Gleason score and aggressiveness in prostate cancer. **npj Digital Medicine**, v. 4, 2021. Disponível em: https://doi.org/10.1038/s41746-021-00469-6. Acesso em: 10 jul. 2025.

OSIECKI, R. *et al.* Prostate cancer morphologies: cribriform pattern and intraductal carcinoma relations to adverse pathological and clinical outcomes—systematic review and meta-analysis. **Cancers**, v. 15, n. 5, p. 1372, 2023. Available from: https://doi.org/10.3390/cancers15051372. Accessed on: 06 may 2025.

OUFATTOLE, Jihane *et al.* Cribriform morphology is associated with higher risk of biochemical recurrence after radical prostatectomy in patients with Grade Group 5 prostate cancer. **Histopathology**, v. 82, n. 7, p. 1089–1097, 2023. Available from: https://doi.org/10.1111/his.14901. Accessed on: 25 jul. 2025.

PARALKAR, D. *et al.* Prostatic adenocarcinoma: molecular underpinnings and treatment-related options. **Urologic Oncology: Seminars and Original Investigations**, v. 42, n. 7, p. 203–210, 2024. Available from: https://doi.org/10.1016/j.urolonc.2024.03.003. Accessed on: 22 jun. 2025.

PARIYAR, Mamta *et al.* Verification and Validation of a Four-Gene Panel as a Prognostic Indicator in Triple Negative Breast Cancer. **Frontiers in Oncology**, v. 12, mar. 2022. Available from: https://doi.org/10.3389/fonc.2022.821334. Accessed on: 19 jul. 2025.

PATASIUS, *et al.*. Prostate cancer screening with psa: Ten years' experience of population based early prostate cancer detection programme in lithuania **Journal of Clinical Medicine**, v. 9, p. 3826, 2020. Available from: https://doi.org/10.3390/jcm9123826. Accessed on: 31 may 2025.

PATON, V. *et al.* Assessing the impact of transcriptomics data analysis pipelines on functional enrichment outcomes. **Nucleic Acids Research**, 2024. Disponível em: https://doi.org/10.1093/nar/gkXXXX. Acesso em: 10 jul. 2025.

PERDANA, Noor R. *et al.* **The risk factors of prostate cancer and its prevention:** a literature review the risk factors of prostate cancer and its prevention. 2016.

PETRANY, Alicia; CHEN, Ruoyu; ZHANG, Shaoqiang; CHEN, Yong. Theoretical framework for the difference of two negative binomial distributions and its application in comparative

analysis of sequencing data. **Genome Research**, v. 34, n. 10, p. 1636–1650, 2024. Available from: https://doi.org/10.1101/gr.278843.123. Accessed on: 19 jul. 2025.

RABILLOUD, T. *et al.* From images to diagnosis: AI in digital pathology for cancer. **Journal of Pathology Informatics**, v. 14, 2023. Available from: https://doi.org/10.4103/jpi.jpi-2223. Accessed on: 10 jul. 2025.

RAWLA, P. *et al.* Epidemiology of prostate cancer: current trends and future perspectives. **Frontiers in Public Health**, v. 10, p. 811044, 2022. Available from: https://doi.org/10.3389/fpubh.2022.811044. Accessed on: 10 jul. 2025.

RICHARDS, Z. *et al.* Prostate cancer risk stratification and decision-making: current status and future directions. **Nature Reviews Urology**, v. 19, p. 45–58, 2022. Available from: https://doi.org/10.1038/s41585-021-00514-0. Accessed on: 10 jul. 2025.

RITCHIE, Matthew E.; PHIPSON, Belinda; WU, Di; HU, Yifang; LAW, Charity W.; SHI, Wei; SMYTH, Gordon K. limma powers differential expression analyses for RNA-sequencing and microarray studies. **Nucleic Acids Research**, v. 43, n. 7, p. e47-e47, 2015. Available from: https://doi.org/10.1093/nar/gkv007. Accessed on: 15 jul. 2025.

ROBINSON, Mark D.; McCARTHY, Davis J.; SMYTH, Gordon K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 2010. Available from: https://doi.org/10.1093/bioinformatics/btp616. Accessed on: 15 jul. 2025.

RUSSO, Giorgio Ivan *et al.* Oncological outcomes of cribriform histology pattern in prostate cancer patients: a systematic review and meta-analysis. **Prostate Cancer and Prostatic Diseases**, v. 26, n. 4, p. 646–654, 2023.

SAEIDI, H. *et al.* Genetic alterations in prostate cancer as diagnostic and prognostic markers. **The Malaysian Journal of Pathology**, v. 45, n. 2, p. 149-155, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/37658525. Accessed on: 12 jul. 2025.

SALACHAN, P. V. *et al.* Spatial whole transcriptome profiling of primary prostate tumors reveals gene expression heterogeneity associated with aggressive disease. **Cancer Research**, 2023. Disponível em: https://doi.org/10.1016/j.canres.2023.01.012. Acesso em: 10 jul. 2025.

SEKHOACHA, Mamello *et al.* Prostate cancer review: genetics, diagnosis, treatment options, and alternative approaches. **Molecules**, v. 27, n. 17, 2022. Available from: https://doi.org/10.3390/molecules27175730. Accessed on: 12 jul. 2025.

SHARMA, M.; MIYAMOTO, H. Percent Gleason pattern 4 in stratifying the prognosis of patients with intermediate-risk prostate cancer. **Translational Andrology and Urology**, v. 7, n. S4, p. S484–S489, 2018. Available from: https://doi.org/10.21037/tau.2018.03.20. Accessed on: 05 may 2025.

SHIMODAIRA, K. *et al.* Significance of the cribriform morphology area ratio for biochemical recurrence in Gleason score 4+4 prostate cancer patients following robot-assisted radical prostatectomy. **Cancer Medicine**, v. 13, n. 5, 2024. Available from: https://doi.org/10.1002/cam4.7086. Accessed on: 13 apr. 2025.

SHORT, E.; WARREN, A. Y.; VARMA, M. Gleason grading of prostate cancer: a pragmatic approach. **Diagnostic Histopathology**, v. 25, n. 10, p. 371–378, out. 2019. Disponível em: https://doi.org/10.1016/j.mpdhp.2019.07.001. Acesso em: 10 jul. 2025.

SINGHAL, N. *et al.* Automated deep learning system for detection and Gleason grading of prostate cancer: performance and agreement with genitourinary pathologists. **Scientific Reports**, v. 12, n. 1, 2022. Disponível em: https://doi.org/10.1038/s41598-022-07217-0. Acesso em: 10 jul. 2025.

SIRINUKUNWATTANA, K. *et al.* Artificial intelligence in pathology: a roadmap for translation. **Nature Medicine**, v. 27, p. 724–732, 2021. Available from: https://doi.org/10.1038/s41591-021-01312-0. Accessed on: 10 jul. 2025.

SIVAGANESH, Vignesh *et al.* Protein Tyrosine Phosphatases: Mechanisms in Cancer. **International Journal of Molecular Sciences**, v. 22, n. 23, p. 12865, 2021. Available from: https://doi.org/10.3390/ijms222312865. Accessed on: 15 jul. 2025.

SONG, H. *et al.* Cribriform architecture in prostate cancer is associated with genomic instability and aggressive features. **European Urology Oncology**, v. 5, n. 3, p. 268–276, 2022. Available from: https://doi.org/10.1016/j.euo.2021.02.006. Accessed on: 10 jul. 2025.

SPRANG, Maximilian *et al.* Overlooked poor-quality patient samples in sequencing data impair reproducibility of published clinically relevant datasets. **Genome Biology**, v. 25, n. 1, p. 222, 2024. Available from: https://doi.org/10.1186/s13059-024-03331-6. Accessed on: 22 jul. 2025.

SPYRATOU, Vasiliki *et al.* Ki67 and prostate specific antigen are prognostic in metastatic hormone naïve prostate cancer. **Acta Oncologica**, v. 62, n. 12, p. 1698-1706, 2023. Available from: https://doi.org/10.1080/0284186X.2023.2254480. Accessed on: 12 jul. 2025.

STOPSACK, K. H. *et al.* Extent, impact, and mitigation of batch effects in tumor biomarker studies using tissue microarrays. **eLife**, v. 10, 2021. Available from: https://doi.org/10.7554/eLife.71265. Accessed on: 11 may 2025.

SUNDARESAN, *et al.*. Prostate-specific antigen screening for prostate cancer: Diagnostic performance, clinical thresholds, and strategies for refinement. **Urologic Oncology: Seminars and Original Investigations** v. 43, p. 41–48, 2025. Available from: https://doi.org/10.1016/j.urolonc.2024.06.003. Accessed on: 11 jun. 2025.

SUNG, Hyuna *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 71, n. 3, p. 209-249, 2021. Available from: https://doi.org/10.3322/caac.21660. Accessed on: 12 jul. 2025.

TABAYOYONG, W.; ABOUASSALY, R. Prostate cancer screening and the associated controversy. **Surgical Clinics of North America**, v. 95, n. 5, p. 1023–1039, 2015. Available from: https://doi.org/10.1016/j.suc.2015.05.001. Accessed on: 12 jul. 2025.

TANG, Haicheng *et al.* The CPNE Family and Their Role in Cancers. **Frontiers in Genetics**, v. 12, jul. 2021. Available from: https://doi.org/10.3389/fgene.2021.689097. Accessed on: 12 jul. 2025.

TENGVALL, Katarina *et al.* Molecular mimicry between Anoctamin 2 and Epstein-Barr virus nuclear antigen 1 associates with multiple sclerosis risk. **Proceedings of the National Academy of Sciences**, v. 116, n. 34, p. 16955–16960, ago. 2019. Available from: https://doi.org/10.1073/pnas.1902623116. Accessed on: 12 jul. 2025.

TOLKACH, Y. *et al.* Artificial intelligence in histopathology: current applications and future perspectives. **Pathology – Research and Practice**, v. 245, 2023. Available from: https://doi.org/10.1016/j.prp.2023.154435. Accessed on: 10 jul. 2025.

TSAI, Harrison K. *et al.* Gene expression signatures of neuroendocrine prostate cancer and primary small cell prostatic carcinoma. **BMC Cancer**, v. 17, n. 1, p. 759, 2017. Available from: https://doi.org/10.1186/s12885-017-3729-z. Accessed on: 12 jul. 2025.

TSAI, Meng-Fong; YU, Shyr-Shen. Distance Metric Based Oversampling Method for Bioinformatics and Performance Evaluation. **Journal of Medical Systems**, v. 40, n. 7, p. 159, 2016. Available from: https://doi.org/10.1007/s10916-016-0516-3. Accessed on: 12 jul. 2025.

ULTSCH, Alfred; LÖTSCH, Jörn. Euclidean distance-optimized data transformation for cluster analysis in biomedical data (EDOtrans). **BMC Bioinformatics**, v. 23, n. 1, p. 233, 2022. Available from: https://doi.org/10.1186/s12859-022-04769-w. Accessed on: 12 jul. 2025.

VAIDYANATHAN, V. *et al.* Prostate cancer: is it a battle lost to age? **Geriatrics**, v. 1, n. 4, p. 27, 2016. Available from: https://doi.org/10.3390/geriatrics1040027. Accessed on: 12 jul. 2025.

VAINIO, Paula *et al.* Phospholipase PLA2G7, associated with aggressive prostate cancer, promotes prostate cancer cell migration and invasion and is inhibited by statins. **Oncotarget**, v. 2, n. 12, p. 1176–1190, 2011. Available from: https://doi.org/10.18632/oncotarget.397. Accessed on: 12 jul. 2025.

VAN DER SLOT, M. A. *et al.* Inter-observer variability of cribriform architecture and intraductal carcinoma of the prostate: relation to Gleason pattern 4 quantification. **Modern Pathology**, v. 33, 2020. Disponível em: https://doi.org/10.1038/s41379-020-0625-x. Acesso em: 10 jul. 2025.

VAN LEENDERS, G. J. L. H. *et al.* The 2019 International Society of Urological Pathology (ISUP) consensus conference on grading of prostatic carcinoma. **American Journal of Surgical Pathology**, v. 44, n. 8, e87–e99, 2020. Disponível em: https://doi.org/10.1097/PAS.0000000000001497. Acesso em: 07 jul. 2025.

VARET, Hugo; BRILLET-GUÉGUEN, Loraine; COPPÉE, Jean-Yves; DILLIES, Marie-Agnès. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. **PLOS ONE**, v. 11, n. 6, p. e0157022, 2016. Available from: https://doi.org/10.1371/journal.pone.0157022. Accessed on: 12 jul. 2025.

VILAÇA, Marta *et al.* Tumor-Infiltrating Lymphocytes in Localized Prostate Cancer: Do They Play an Important Role? **Cureus**, v. 15, n. 1, p. e34007, 2023. Available from: https://doi.org/10.7759/cureus.34007. Accessed on: 12 jul. 2025.

VLAJNIC, T.; EPPENBERGER-CASTORI, S.; BUBENDORF, L. Protocols for tissue microarrays in prostate cancer studies.**Methods in Molecular Biology**. Humana Press, 2018. v.

1786, p. 103–116. Available from: https://doi.org/10.1007/978-1-4939-7845-8_6. Accessed on: 12 jul. 2025.

WALKER, Simon R. *et al.* Decreased ATM protein expression is substantiated with PTEN loss in defining aggressive phenotype of prostate cancer associated with lethal disease. **European Urology Open Science**, v. 29, p. 93-101, 2021. Available from: https://doi.org/10.1016/j.euros.2021.05.004. Accessed on: 12 jul. 2025.

WANG, Y. *et al.* Integrated analysis of transcriptomic and genomic alterations in prostate cancer reveals potential prognostic markers. **Frontiers in Oncology**, v. 12, p. 881246, 2022. Available from: https://doi.org/10.3389/fonc.2022.881246. Accessed on: 2 aug. 2025.

WONG, Hong Yuen *et al.* Single cell analysis of cribriform prostate cancer reveals cell intrinsic and tumor microenvironmental pathways of aggressive disease. **Nature Communications**, v. 13, n. 1, p. 6036, 2022. Available from: https://doi.org/10.1038/s41467-022-33780-1. Accessed on: 22 jul. 2025.

XIA, *et al.*. Surgical delay and pathological outcomes for clinically localized high-risk prostate cancer **JAMA Network Open**, v. 3, p. 2028320, 2020. Available from: https://doi.org/10.1001/jamanetworkopen.2020.28320. Accessed on: 03 jul. 2025.

XIE, Jun *et al.* Combination of theoretical analysis and experiments: Exploring the role of PLA2G7 in human cancers, including renal cancer. **Heliyon**, v. 10, n. 6, p. e27906, 2024. Available from: https://doi.org/10.1016/j.heliyon.2024.e27906. Accessed on: 22 jul. 2025.

YANG, Kun; LI, Jianzhong; GAO, Hong. The impact of sample imbalance on identifying differentially expressed genes. **BMC Bioinformatics**, v. 7, supl. 4, p. S8, 2006. Available from: https://doi.org/10.1186/1471-2105-7-S4-S8. Accessed on: 18 jul. 2025.

YANG, M.; WU, J.; ZHANG, Y. Prostate cancer incidence and mortality around the globe: An analysis of the Global Cancer Observatory data. **Social Sciences**, v. 13, n. 8, p. 705–720, 2022. Available from: https://doi.org/10.4236/ss.2022.138047. Accessed on: 2 aug. 2025.

YANG, Meng-Ling *et al.* SLC13A4 Might Serve as a Prognostic Biomarker and be Correlated with Immune Infiltration into Head and Neck Squamous Cell Carcinoma. **Pathology and Oncology Research**, v. 27, nov. 2021. Available from: https://doi.org/10.3389/pore.2021.1609967. Accessed on: 18 jul. 2025.

YANG, Min-Hui *et al.* Nuclear lncRNA HOXD-AS1 suppresses colorectal carcinoma growth and metastasis via inhibiting HOXD3-induced integrin beta3 transcriptional activating and MAPK/AKT signalling. **Molecular Cancer**, v. 18, n. 1, p. 31, dez. 2019. Available from: https://doi.org/10.1186/s12943-019-0955-9. Accessed on: 18 jul. 2025.

YE, Junhui *et al.* MicroRNA-671-5p inhibits cell proliferation, migration and invasion in non-small cell lung cancer by targeting MFAP3L. **Molecular Medicine Reports**, v. 29, n. 6, p. 101, abr. 2024. Available from: https://doi.org/10.3892/mmr.2024.13225. Accessed on: 05 jul. 2025.

YE, Min *et al.* Heterogeneous programmed death-ligand 1 expression in gastric cancer: comparison of tissue microarrays and whole sections. **Cancer Cell International**, v. 20, n. 1, p.

186, 2020. Available from: https://doi.org/10.1186/s12935-020-01273-0. Accessed on: 12 jul. 2025.

YOU, Yue *et al.* Modeling group heteroscedasticity in single-cell RNA-seq pseudo-bulk data. **Genome Biology**, v. 24, n. 1, p. 107, 2023. Available from: https://doi.org/10.1186/s13059-023-02949-2. Accessed on: 05 jul. 2025.

ZENGIN, Zeynep B. *et al.* Clinical implications of AR alterations in advanced prostate cancer: a multi-institutional collaboration. **Prostate Cancer and Prostatic Diseases**, fev. 2024. Available from: https://doi.org/10.1038/s41391-024-00805-3. Accessed on: 05 jul. 2025.

ZHANG, L. *et al.* Cribriform morphology is an independent predictor of adverse outcome in prostate cancer: molecular correlates and clinical validation. **Journal of Pathology**, v. 257, n. 3, p. 312–323, 2022. Available from: https://doi.org/10.1002/path.5832. Accessed on: 10 jul. 2025.

ZHANG, Y. *et al.* AI-powered digital pathology predictive modeling for prostatic cribriform patterns and aggressiveness. **npj Digital Medicine**, v. 6, 2023. Available from: https://doi.org/10.1038/s41591-023-02614-9. Accessed on: 10 jul. 2025.

ZHANG, Zhao-Cun *et al.* Tripartite motif–containing 9 promoted proliferation and migration of bladder cancer cells through CEACAM6-Smad2/3 axis. **Journal of Cell Communication and Signaling**, v. 17, n. 4, p. 1323–1333, dez. 2023. Available from: https://doi.org/10.1007/s12079-023-00766-7. Accessed on: 05 jul. 2025.

ZHENG, Weili *et al.* Identification of PLA2G7 as a novel biomarker of diffuse large B cell lymphoma. **BMC Cancer**, v. 21, n. 1, p. 927, 2021. Available from: https://doi.org/10.1186/s12885-021-08660-4. Accessed on: 05 jul. 2025.