

# Utility of Three-Protein Panels in the Separation of Aggressive Prostate Cancer from Non-Aggressive Tumors

Qing Kay Li (✉ [qli23@jhmi.edu](mailto:qli23@jhmi.edu))

Johns Hopkins Hospital

Tung-shing Mamie Lih

Johns Hopkins Hospital

Yuefan Wang

Johns Hopkins Hospital

Yingwei Hu

Johns Hopkins Hospital

Naseruddin Hoti

Johns Hopkins Hospital

Daniel Chan

Johns Hopkins Hospital

Hui Zhang

Johns Hopkins Hospital

---

## Research Article

**Keywords:** prostate cancer, protein biomarkers, aggressive prostate cancer, immunohistochemistry, tumor tissue microarray

**Posted Date:** November 15th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1047551/v1>

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

[Read Full License](#)

---

# Abstract

## Background

Prostate cancer (PCa) is a heterogeneous group of tumors, including non-aggressive (NAG) and aggressive (AG) subtypes, with variable clinical outcomes. We assessed the diagnostic utility of selected protein markers to identify AG tumors.

## Methods

The TMA was constructed, including NAG and AG. 12 protein markers were evaluated using the TMA by IHC stains. The makers were also evaluated for their potential utility as single or panels for distinguishing AG from NAG tumors.

## Results

The higher expressions of four protein markers, including prostate specific membrane antigen (PSMA), phospho-EGFR, androgen receptor (AR), and P16, were identified in AG tumors of Gleason score 4 and 5. In contrast, Galectin-3, DPP4 and MAN1B1 revealed stronger staining patterns in NAG tumors. Sensitivity and specificity of individual marker varied widely. In tow-marker panels, especially in the panel of DPP4 and PSMA, the specificity was 38.46% at 95% sensitivity. To further improve the detection ability, we combined DPP4 and PSMA with either Galectin-3 or phospho-EGFR into three-marker panels. The specificity achieved >46% at 95% sensitivity and AUC was >0.85.

## Conclusions

Our panels can be used to improve the separation of AG from NAG tumor and to add in the optimization of the treatment strategy for patients.

## Introduction

Prostate cancer (PCa) is a heterogeneous group of tumors with variable clinical outcomes<sup>1-3</sup>. Most PCa presents as a localized disease with low or no risk for the tumor progression, and patients can be managed by the active surveillance program rather than aggressive/surgical interventions. Only a subset of PCa has an aggressive behavior leading to tumor progression, metastasis and cancer-related death<sup>4-7</sup>. Overdiagnosis and overtreatment of indolent PCa are still clinical issues in the management of PCa patients. Several recent studies, including the US Preventive Services Task Force (USPSTF), have shown that the incidence of localized disease continues to decline, whereas, the incidence of advanced-stage disease continues to arise in men over 50-year old<sup>4-7</sup>. Multiple risk stratification systems have been

developed to separate the high-risk aggressive PCa (AG) from low-risk non-aggressive indolent tumors (NAG), including the combination of both clinical and pathological parameters (such as serum PSA levels, Gleason score, ISUP grade of the tumor, as well as clinical and pathological staging). However, these tools are still insufficient for the prediction of disease progression or separation of AG from NAG<sup>4,6-8</sup>. Therefore, there is an urgent clinical need to identify AG in order to optimize the treatment strategy for patients.

Current genomic studies have identified a spectrum of molecular abnormalities associated with PCa<sup>9-13</sup>. These studies demonstrated a distinct molecular abnormalities in subtypes of PCa. In the Cancer Genome Atlas (TCGA) study, the comprehensive analysis of a large cohort revealed that 74% of 333 primary PCAs fell into one of seven subtypes defined by specific gene fusions (*ERG*, *ETV1/4*, *FLI1*) or mutations (*SPOP*, *FOXA1*, *IDH1*)<sup>9</sup>. Among numerous genomic aberrations, *SPOP* and *FOXA1* mutants had the highest levels of AR-induced transcripts<sup>9</sup>. The study of whole-genomes and tumor methylomes revealed several unique tumor-specific RNA and methylation patterns in AG tumors<sup>10</sup>. In addition, many other epigenetic alterations have been also identified in PCa, involving in EGFR, PI3K, MAPK signaling pathways and DNA repair genes<sup>11,12</sup>, loss of *PTEN* and alterations in *TMPRSS2-ERG* fusions<sup>12-14</sup>. These studies reveal not only genomic heterogeneity among primary PCa, but also identify potentially actionable therapeutic targets.

Furthermore, the dysregulation of genomics in PCa leads to aberrant transcription and expression of cellular proteins. Based upon multi-omic studies of proteogenomics and comprehensive proteomic analysis of tumor tissue, many potential protein biomarkers and molecular mechanisms associated with AG tumors have also been suggested<sup>15-23</sup>. Sinha A, et al. studied the proteomic signatures of 76 localized, intermediate-risk PCa tumor tissues, and identified four distinct protein clusters correlated with five clinical groups as well as genomic subtypes of PCa<sup>15</sup>. Interestingly, they also found that changes of mRNA abundance could not reflex the protein abundance variability, indicating the importance of proteomic study of tumors. In a comparative study of PCa cell lines with PCa tumor tissues, 12 mutant peptides were identified to be differentially expressed in PCa<sup>16</sup>. Our previous proteomic studies identified certain protein biomarkers were differentially expressed, including 14 high-expression proteins in AG tumors and 14 high-expression proteins in NAG tumors<sup>17</sup>. Similarly, several other studies were also demonstrated that proteins were differentially expressed in subtypes of PCa<sup>18-23</sup>. In the study of 28 primary PCa with Gleason scores ranged from 6 to 9, authors found an increased expression of pro-NPY which was associated with a poor prognosis<sup>18</sup>. Recently, we identified a decreased proteinase activity in AG tumors<sup>19</sup>. All these findings demonstrate an extensive involvement of intracellular proteins and signaling pathways in PCa. Therefore, further evaluation of the roles of potential protein markers as independent predictors of pathological AG tumor is necessary<sup>20-23</sup> in order to optimize the clinical management.

In PCa, the Gleason score of a tumor has been considered to be an indicator of the aggressiveness of the disease, and is assigned based upon the morphology of the dominant tumor pattern plus secondary

tumor pattern<sup>24,25</sup>. Based upon the Gleason score of tumors, a PCa grading system was further developed by the International Society of Urological Pathology (ISUP) to predict the clinical behavior of the tumor<sup>24</sup>. In ISUP grading system, grade group 1 (Gleason score 6 or less) has a low risk for disease progression, grade group 2 (Gleason score 3+4 = 7) and grade group 3 (Gleason score 4+3 = 7) have intermediate risk for progression, grade group 4 (Gleason score 8) has a high risk for progression, and grade group 5 (Gleason score 9 or 10) has the highest risk for tumor progression<sup>24</sup>. Although both ISUP grade group 2 and 3 are considered to be intermediate risk groups, studies have shown that grade group 2 has a favorable prognosis than that of grade group 3<sup>25,26</sup>. Thus, the current guideline suggests that the active surveillance program should be selected for grade group 2 and surgical interventions should be selected for grade group 3 or higher PCa<sup>27,28</sup>. In patients, a small amount of tumor tissue often obtained for the initial evaluation of tumor Gleason scores. However, the procedure can be difficult and may not represent the true nature of the tumor. For example, the Gleason score, particularly the secondary pattern of an aggressive tumor may be difficult to identify and/or assigned in certain cases due to the minute focus of the representative morphology<sup>24,29</sup>. Studies have shown that patients with biopsy proven grade group 1 and grade group 2 could upgrade to higher grade groups after evaluation of the surgical resected whole tumor tissue<sup>29</sup>. Therefore, an integrated evaluation of the tumor morphology in combination with the expression of protein markers is necessary to improve the identification of AG from NAG tumors.

To assess the utility of potential protein markers and to understand molecular features of tumor progression in PCa, we evaluated expressional patterns of 12 protein biomarkers identified by proteomic and genomic studies. The immunohistochemistry (IHC) was performed by using PCa tumor microarrays (TMA), including both indolent NAG and AG subtypes and tumor-matched normal/benign adjacent tissue (NAT). The purposes of our study were to evaluate the utility of PCa-associated proteins, to optimize the performance of individual marker by combining them into panels, and to assess the protein biomarker panels, which may have a potential for the development of clinical assay to separate AG from NAG PCa.

## Results

### Clinical Information

In our cohort, the median age of patients was 61 years, ranging from 40 to 73 years. Based upon the grading criteria of International Society of Urological Pathology (ISUP) and the morphological feature of the dominant nodule, the ISUP grade of our cohort were: 20 cases of Grade 1, 18 cases of Grade 3, 5 cases of Grade 4, and 14 cases of Grade 5. In addition, 7 cases of Grade 2 were found in tertiary nodules. The pathological stages were 30 cases of pT2, 1 case of pT2x, 13 cases of pT3A, 11 cases of pT3B and 2 cases of pT4 (**Table 2, Supplementary Table 1**). Taken together, 20 indolent NAGs with the Gleason score of 6, and 37 AGs with the Gleason score  $\geq 7$  were included in the TMA.

Overall, our TMA contains a total of 218 tumor cores, including 92 cores (42.2%) of Gleason score 3, 63 cores (28.9%) of Gleason score 4, and 63 cores (28.9%) of Gleason score 5; and 113 cores of NAT.

## IHC Staining Pattern of Individual Protein in PCa TMA

IHC stains of 12 proteins were performed on the TMA. The information of primary antibodies is summarized in **Table 1**. Among them, antibodies for 7 proteins presented variably staining patterns (**Figure 1A**). The majority of protein markers revealed membrane and cytoplasmic staining patterns, except P16, which revealed nuclear staining pattern. The staining patterns of each protein in tumors and NAT were analyzed using a semi-quantitative scoring system.

The IHC staining patterns of antibodies to 7 proteins are summarized in **Table 3**. The higher expressions of four proteins, including PSMA, phospho-EGFR, androgen receptor (AR), and P16, were identified in AG tumors. In contrast, three antibodies, including anti-Galectin-3, anti-DPP4, and anti-MAN1B1, revealed stronger staining patterns in Gleason score 3 tumors, but weak staining patterns in Gleason score 4 and Gleason 5 tumors. IHC score 1, 2 and 3 were used as cut-off score for positive or negative stain in tumors, respectively. Both PSMA and phospho-EGFR had a positive correlation with Gleason scores of the tumor, whereas Galectin-3 and DPP4 had negative correlation with Gleason scores of the tumor (**Figure 1B**). The correlations of IHC scores of PSMA, phospho-EGFR, Galectin-3 and DPP4 with Gleason scores of tumors are shown in **Figure 2**.

We did not detect the expression of total EGFR with antibody D38B1, PD-1 with antibody NAT105, PD-L1 with antibodies 22C3 and SP142, and PTEN with antibody 6H2.1 in PCa.

## The Sensitivity and Specificity of Individual Proteins

Based upon the individual staining pattern, the receiver operating characteristic (ROC) analysis was performed. The value of an area under the curve (AUC) of individual marker was compared. The sensitivity and specificity of phospho-EGFR, Galectin-3, DPP4 and PSMA for distinguishing AG from NAG are summarized in **Figure 3**.

Among individual marker, values of AUC ranged from 0.48 to 0.7, with best performance of DPP4 and PSMA. Expressions of DPP4 and PSMA were significantly altered in tumors with the Gleason  $\geq 4$  ( $p < 0.05$ ) and demonstrated a better performance than phospho-EGFR and galectin-3 in the separation of AG tumors from NAG tumors (**Figure 3A**). To directly compare the performance of individual marker, we fixed the sensitivity at 95% and then compared the specificity. At 95% sensitivity, the specificity was ranged from 0% to 8.79% (**Figure 3B**). By examining the individual marker at its best cutoff point on ROC curves (the maximal summed sensitivity and specificity), we found that phospho-EGFR, Galectin-3, DPP4 and PSMA, had specificities of 49.5%, 89%, 73.6% and 89%; and the corresponding sensitivities were 68.1%, 36.3%, 79.6%, and 68.1%, respectively (**Figure 3B**). Among these markers, Galectin-3 and PSMA had the best specificity of 89%; and DPP4 had the best sensitivity of 79.6%. To evaluate statistical stability of the performance of markers for AG tumor detection, we used both label permutation and bootstrap methods (**Figure 3C**). Again, both DPP4 and PSMA demonstrated higher stability than that of phospho-EGFR and Galectin-3. Taken together, the reduced expression of DPP4 and elevated expression of PSMA could be used as a signature of aggressiveness of PCa.

## Further Construction and Evaluation of Protein Panels in Separation of AG Tumors

Based upon the performance of individual protein marker, we combined individual protein biomarker into two- and three-marker panels, and evaluated their performances in the separation of AG from NAG tumors.

The two-marker panels were constructed by using combinations of Galectin-3 plus PSMA, phospho-EGFR plus PSMA, DPP4 plus Galectin-3, and DPP4 plus PSMA (**Figure 4**). The overall performance for distinguishing AG and NAG tumors was improved when using panels, compared to using individual maker. All AUCs of these panels were  $> 0.70$  (**Figure 4A**). In panels, specificities were  $> 68\%$  (85.7%, 86.8%, 68.1% and 76.9%); and sensitivities were  $> 69\%$  (71.7%, 69%, 87.6% and 85%) (**Figure 4B**). The label permutation and bootstrap analyses demonstrated the stable performance of two-marker panels (**Figure 4C**). Additionally, the panel of DPP4 and PSMA showed a better specificity (38.46%) when fixed sensitivity at 95% compared to individual markers and other two-marker panels (**Figure 4B**).

To further improve the specificity of these markers using a 95% sensitivity as cutoff value, we also constructed three-marker panels and evaluated their performance. These panels included combinations of DPP4 plus Galectin-3 plus phospho-EGFR, DPP4 plus Galectin-3 plus PSMA, and DPP4 plus phospho-EGFR plus PSMA (**Figure 5**). All AUCs were further improved to  $> 80\%$  (**Figure 5A** and **Figure B**). Specificities and sensitivities of three-marker panels were as follows: 75.8% and 83.2% in the panel of DPP4 plus Galectin-3 plus phospho-EGFR, 83.5% and 76.1% in the panel of DPP4 plus Galectin-3 plus PSMA, 81.3% and 79.6% in the panel of DPP4 plus phospho-EGFR plus PSMA (**Figure 5B**). All specificities of three-marker panels were  $> 75\%$  (75.8%, 83.5%, and 81.3%); and sensitivities were  $> 76\%$  (83.2%, 76.1%, 79.6%). The random models (label permutation analysis) and the real data were well-separated, indicating that the performances of three-marker panels were reliable (**Figure 5C**). We observed that three-marker panels had much better performance than that of individual makers as well as two-marker panels, and the specificity at 95% sensitivity was improved, especially in panels composed of both DPP4 and PSMA. Specificities at 95% sensitivity of these panels reached 48.35% and 46.15%, respectively (**Figure 5B**).

Taken together, our data demonstrated that three-marker panel containing DPP4 and PSMA can significantly improve the separation of AG from NAG, in the comparison with individual marker or two-marker panels.

## Discussion

Prostate cancer (PCa) is the most common cancer and the second leading cause of cancer death in men in the United States; with estimated new cases and cancer-related deaths in 2020 were 192,000 and 33,000, respectively<sup>1-4</sup>. The majority of PCa presents as a indolent tumor, in which the patient can be observed in an active surveillance program<sup>4,24-28</sup>. Only about 10% of PCa presents as an aggressive disease with high risk of tumor progression<sup>4,26-28</sup>. Therefore, in order to accurately predict the high-risk

tumor and to limit overtreatment of indolent tumor, it is crucial to distinguish AG PCa from NAG tumors. Although both Gleason score and newly developed ISUP grade group system have been used to assess the potential clinical behavior of the tumor, these systems still have certain limitations to guide the therapeutic decision for patients<sup>24</sup>.

To better understand the molecular mechanism and to identify AG tumors, great efforts have been established to profile the proteogenomic landscape of PCa<sup>9-23</sup>. Our previous large-scale quantitative proteomic studies of PCa tumor tissue identified a spectrum of cellular proteins to be up- or down-regulated in subtypes of PCa<sup>8,17,19</sup>. These differentially expressed tissue proteins identified by MS approach are particularly interesting as biomarker candidates because of the high likelihood of their detectability in tumor tissue<sup>17,19</sup>. Using an immunochemical approach, we assessed the potential detectability of selected candidate protein markers in tumor tissue, and their utility as individual markers and/or as protein panels in the separation of AG tumors.

In this study of 12 markers, expressional patterns of four proteins, including PSMA, phospho-EGFR, AR, and MAN1B1, demonstrated higher levels in AG tumors in comparison to NAG. In contrast, three protein markers, including Galectin-3, DPP4 and P16, revealed stronger staining patterns in Gleason score 3 tumors, but weak staining patterns in Gleason score 4 and Gleason 5 tumors. The AUCs of phospho-EGFR, Galectin-3, DPP4 and PSMA were 0.48, 0.54, 0.68, and 0.7, respectively. Among the individual marker, PSMA showed the best discrimination power with the specificity of 89%, where DPP4 showed a sensitivity of 79.6%. These two markers demonstrated a better performance than phospho-EGFR and Galectin-3 in AG tumor detection. However, these two markers had limitations when we considered the specificity at 95% sensitivity. Nevertheless, reduced expression of DPP4 and elevated expression of PSMA could be used together as signatures of aggressiveness of PCa.

PSMA is a type II transmembrane glycoprotein, containing 750-amino acid. It has a long C-terminal extracellular domain and a short N-terminal intracellular domain<sup>31</sup>. Its extracellular domain has enzymatic activity functioning as folate hydrolase I or glutamate carboxypeptidase II<sup>32</sup>. In benign prostate tissue, the expression of PSMA is low. This low level of expressional pattern is also identified in the kidney, small intestine, and brain tissue. However, the expression of PSMA is significantly increased in PCa; and the overexpression of PSMA is also correlated with the disease progression and the tumor metastasis in PCa<sup>33,34</sup>. Similarly, DPP4 is also a type II transmembrane glycoprotein, but it has the serine exopeptidase activity. DPP4 plays a critical role in regulating cellular proliferation and migration<sup>35</sup>. The aberrant oncogenic and tumor suppressor activity of DPP4 have been identified in cancers<sup>36,37</sup>. A reduced serum DPP4 level is also found in PCa patients, especially in patients with metastatic disease<sup>38</sup>. In the study of primary and metastatic PCa, we recently identified that the decreased DPP4 expression and activity is associated with PCa aggressiveness<sup>17,19</sup>. The findings of decreased DPP4 levels in aggressive and metastatic PCa suggest its critical role in AG tumors. Galectin-3 is a member of the lectin superfamily and plays critical roles in regulating cellular signaling pathways and cancer progression<sup>39</sup>. In prostate cancer, its expression elevated in the early stages of tumors, but this expression gradually

decreased over disease progression and was completely lost in advanced stage tumors<sup>39-41</sup>. In the metastatic PCa, Galectin-3 regulates tumor cells to form aggregates and adhere to the microvascular endothelium<sup>42</sup>. Based upon its expression and biological roles, Galectin-3 has been suggested to be a predictive marker for the biochemical recurrence of PCa<sup>40,41</sup>. EGFR is a transmembrane glycoprotein and activated by the dimerization upon a ligand binding<sup>43</sup>. The phosphorylation of EGFR leads to several downstream intracellular phosphorylations<sup>43-46</sup>. EGFR activation plays a key role in cell survival, proliferation, migration and differentiation, including in PCa progression<sup>45,46</sup>.

Based upon the performance of individual protein markers and the important biological functions of these protein markers, particularly DPP4 and PSMA, we combined individual biomarker into several panels and evaluated their performances in the separation of AG from NAG tumors. The two-marker panels were constructed by using combinations of PSMA or DPP4 with either Galectin-3 or phospho-EGFR. Higher AUCs were achieved using two-marker panels compared to individual marker, and indicated a significant improvement in differentiating AG tumors and NAG tumors. All specificities of two-marker panels were over 68% (85.7%, 86.8%, 68.1% and 76.9%); and all the sensitivities were > 69% (71.7%, 69%, 87.6% and 85%). Furthermore, the specificity at 95% sensitivity was improved in general by using two-marker panels. Among the combinations, two-marker panel consisted of DPP4 and PSMA demonstrated the highest specificity (38.46%) at 95% sensitivity, indicating the best performance.

To investigate whether a higher discrimination power could be further achieved, three-marker panels were also constructed by using combinations of both DPP4 and PSMA with either Galectin-3 or phospho-EGFR. In three-marker panels, all AUCs were improved to >0.83, indicating a further improvement in the separation of AG from NAG. All specificities of three-marker panels were > 75% (75.8%, 83.5%, and 81.3%); and all sensitivities were > 76% (83.2%, 76.1%, and 79.6%). In three-marker panels containing both DPP4 and PSMA, specificities at 95% sensitivity reached 46.15% and 48.35%, respectively. These three-marker panels demonstrated the best performance, compared to individual marker and other two- and three-marker panels.

The unique feature of our study is the integrative analysis of protein expressions with a spectrum of tumor Gleason scores, including Gleason score 6 NAG (Gleason score 3+3), Gleason score  $\geq 7$  AG (Gleason score 3+4, 4+3, 4+4, 4+5, 5+4 and 5+5) and NAT. IHC stains of these protein markers correlated and validated previously studies of differential expression of proteins in subtypes of PCa. This integrative IHC study with the previously defined aggressive PCa subtypes demonstrated that three-marker panels can be used for the separation of aggressive PCa from indolent tumors. Our three-marker panel can be served as a signature of AG tumors.

Furthermore, our findings also demonstrated that loss of Galectin-3 expression and DPP4 activity may promote prostate cancer aggressiveness. The consequence of the decrease expression of these two proteins and subsequent increase in bio-active of phospho-EGFR and PSMA promote tumor cell proliferation and disease progression. It also helps us to gain further knowledge into the proteomic

heterogeneity of aggressive PCa and to investigate the molecular taxonomy of the tumor for future diagnostic, prognostic, and therapeutic stratification.

In summary, we assessed the diagnostic value of selected protein markers in the identification of AG tumors using TMA and IHC. The higher expressions of four protein markers, including PSMA, phospho-EGFR, AR, and P16, were identified in AG tumors of Gleason score 4 and 5. In contrast, three protein markers, including Galectin-3, DPP4 and MAN1B1, revealed stronger staining patterns in NAG tumors of Gleason score 3. The sensitivity and specificity of individual marker for distinguishing AG were variable and relatively low. We constructed two- and three-marker panels. The combination of two tissue markers could provide better separation of AG from NAG tumors, especially the panel composed of DPP4 and PSMA. We observed further improvement when combining DPP4 and PSMA with Galectin-3 (AUC of 0.85) as well as combining DPP4 and PSMA with phospho-EGFR (AUC of 0.86). More importantly, higher specificities of 46.1% and 48.35% were achieved by using the aforementioned three-marker panels at fixed 95% sensitivity. These panels can be used to assess the aggressiveness of PCa and to improve the separation of AG from NAG tumor using tumor tissue. The utility of these panels provides an additional diagnostic tool to address the urgent clinical need and to optimize the treatment strategy for PCa patients.

## Methods

### Case Selection

PCa cases were collected from radical prostatectomy (except one case was collected from transurethral resection of the prostate (TURP)) with informed consents and in a manner to protect patients' identity. A total of 57 cases were included in the study, including Gleason score 3 (i.e. 3+3), 4 (i.e. 3+4, 4+3, or 4+4), or 5 (i.e. 5+4 or 4+5) tumors. Among them, 54 cases were collected between January 2002 and December 2009, and additional 3 cases were collected in 2012. Electronic medical records were reviewed and the clinical and pathological data, including age, TNM T-stage, N-stage, M-stage, were obtained (**Supplementary Table 1**). The pathological stages of PCa were determined according to the eighth edition of the AJCC guidelines<sup>30</sup>. AG and NAG tumor were defined using the criteria of the International Society of Urological Pathology (ISUP)<sup>24</sup>.

The study was approved by the Institutional Review Board of Johns Hopkins Medical Institutions. In addition, all methods performed in the study are in accordance with the relevant guidelines and regulations.

### Construction of Tissue Microarray

The PCa tissue microarray (TMA) was constructed using above surgical resected tumors (n=57 cases). All tumor tissue blocks were fixed in 10% formalin and embedded in paraffin. In addition to the original pathology reports, the hematoxylin and eosin (H&E) stained tumor sections were re-reviewed by the American Board of Pathology certified pathologist (QKL) prior to TMA construction to ensure the

representation of tumor area and adjacent tumor-matched benign tissue (NAT). Cores (in diameter of 0.6 mm) were obtained from tumor tissue and/or NAT on each representative paraffin block and transplanted into the recipient TMA block. Of 57 cases, 215 cores of PCa and 111 cores of NAT were included in the TMA.

## Resources of Primary Antibodies

We evaluated 12 proteins in the study. Details of primary antibodies are summarized in **Table 1**.

## Immunohistochemical Staining of PCa TMA

Immunohistochemical (IHC) study of protein biomarkers was performed on TMA using a Ventana Discovery Ultra autostainer (Roche Diagnostics). TMA was cut into 4- $\mu$ m sections prior to IHC study. Briefly, following dewaxing and rehydration on board, epitope retrieval was performed using Ventana Ultra CC1 buffer (catalog# 6414575001, Roche Diagnostics) at 96°C for 48 minutes. Primary antibodies were applied at 36°C for 60 minutes. Primary antibodies were detected using an anti-mouse and/or anti-rabbit HQ detection system (catalog# 7017936001 and 7017812001, Roche Diagnostics) followed by Chromomap DAB IHC detection kit (catalog # 5266645001, Roche Diagnostics), counterstaining with Mayer's hematoxylin, dehydration and mounting.

Distinct membranous, cytoplasmic or nuclear staining was considered in each protein IHC staining. The intensity of IHC staining pattern on each protein was semi-quantitatively by two researchers QKL (the American Board of Pathology certified pathologist) and NH, using a 4-tier system as: 0 (0%, no staining), 1 (<10%, weak and focally staining), 2 (10-50%, medium and focally staining), or 3 (>50%, strong and diffusely staining) in tumor cells. All IHC stains were scanned using Concentriq (Proscia Inc, Philadelphia, PA <https://proscia.com>) and stored as digital files. Depending on the TMA section, some of cores could not be evaluated due to the loss of tissue cores during the process (please see individual staining pattern in the result).

## Statistical Analysis

The discriminatory power of each protein marker panel (composed of  $\geq 1$  candidate marker), using logistic regression, was evaluated via receiver operating characteristic (ROC) curve analysis. To ensure statistical stability of the results, we used bootstrap resampling (n=500) of the data to construct and evaluate the predictive models of protein marker panels. Bootstrap resampling with label permutation was also carried out to generate random models for examining the reliability of the panels. The mean ROC curves were depicted based on bootstrap resampling results and an area under the curve (AUC) was computed for each mean ROC curve. All the analyses were carried out in R (version 3.5). The predictive models were built using caret (version 6.0-85) and ROC curves were generated using pROC (version 1.13).

The difference of IHC staining in AG, NAG and NAT was assessed using the Students' t test. The *p*-values <0.05 were considered statistically significant.

## Abbreviations

PCa: prostate cancer; AG: aggressive prostate cancer; NAG: non-aggressive prostate cancer; C: control; USPSTF: US preventive services task force; ISUP: international society of urological pathology MS: mass spectrometry; TMA: tumor microarray; TCGA: the Cancer Genome Atlas; H&E: hematoxylin and eosin; ACN: Acetonitrile; CV: coefficient of variations; SD: standard deviation; HPA: Human Proteome Atlas; QC: quality control; FDA: Food and Drug Administration. IHC: immunochemistry; ROC: receiver operating characteristic curve; AUC: area under curve; AR: androgen receptor. pEGFR: phospho-EGFR. PSMA: prostate specific membrane antigen.

## Declarations

### Acknowledgments

This work is supported in part by the National Institutes of Health under grants and contracts of National Cancer Institute, the Early Detection Research Network (EDRN, U01CA152813) and Clinical Proteomics Tumor Analysis Consortium (CPTAC, U24CA210985).

### Author Contributions

Q.K.L. H.Z. contributed to experimental design, organization of data and drafting the manuscript; QKL, HZ, YFW, TSL, NH, YWH. contributed to data analysis. DWC contributed to develop project and critics.

### Additional Information

#### Competing Interesting Statement (Conflict of Interests)

All authors declare that they have no competing interests.

### Supporting Information

Supporting information includes Tables S1 Clinical information (.xlsx).

### Consent for publication

This manuscript has been read and approved by all the authors to publish and is not submitted or under consideration for publication elsewhere.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians* 2020; **70**(1): 7-30.
2. Hu JC, Nguyen P, Mao J, Halpern J, Shoag J, Wright JD *et al.* Increase in Prostate Cancer Distant Metastases at Diagnosis in the United States. *JAMA Oncology* 2017; **3**(5): 705-707.
3. Jemal A, Culp MB, Ma J, Islami F, Fedewa SA. Prostate Cancer Incidence 5 Years After US Preventive Services Task Force Recommendations Against Screening. *J Natl Cancer Inst* 2021; **113**(1): 64-71.
4. Force USPST. Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2018; **319**(18): 1901-1913.
5. Negoita S, Feuer EJ, Mariotto A, Cronin KA, Petkov VI, Hussey SK *et al.* Annual Report to the Nation on the Status of Cancer, part II: Recent changes in prostate cancer trends and disease characteristics. *Cancer* 2018; **124**(13): 2801-2814.
6. Houston KA, King J, Li J, Jemal A. Trends in Prostate Cancer Incidence Rates and Prevalence of Prostate Specific Antigen Screening by Socioeconomic Status and Regions in the United States, 2004 to 2013. *J Urol* 2018; **199**(3): 676-682.
7. Islami F, Siegel RL, Jemal A. The changing landscape of cancer in the USA – opportunities for advancing prevention and treatment. *Nature Reviews Clinical Oncology* 2020; **17**(10): 631-649.
8. Li QK, Chen L, Ao M-H, Chiu JH, Zhang Z, Zhang H *et al.* Serum fucosylated prostate-specific antigen (PSA) improves the differentiation of aggressive from non-aggressive prostate cancers. *Theranostics* 2015; **5**(3): 267-276.
9. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* 2015; **163**(4): 1011-1025.
10. Houlahan KE, Shiah Y-J, Gusev A, Yuan J, Ahmed M, Shetty A *et al.* Genome-wide germline correlates of the epigenetic landscape of prostate cancer. *Nature Medicine* 2019; **25**(10): 1615-1626.
11. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014; **505**(7484): 495-501.
12. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* 2010; **18**(1): 11-22.
13. Grasso CS, Wu Y-M, Robinson DR, Cao X, Dhanasekaran SM, Khan AP *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; **487**(7406): 239-243.
14. Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, Sivachenko AY *et al.* The genomic complexity of primary human prostate cancer. *Nature* 2011; **470**(7333): 214-220.
15. Sinha A, Huang V, Livingstone J, Wang J, Fox NS, Kurganovs N *et al.* The Proteogenomic Landscape of Curable Prostate Cancer. *Cancer Cell* 2019; **35**(3): 414-427.e416.
16. Kwon OK, Ha Y-S, Lee JN, Kim S, Lee H, Chun SY *et al.* Comparative Proteome Profiling and Mutant Protein Identification in Metastatic Prostate Cancer Cells by Quantitative Mass Spectrometry-based Proteogenomics. *Cancer Genomics Proteomics* 2019; **16**(4): 273-286.

17. Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B *et al.* Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. *Mol Cell Proteomics* 2014; **13**(7): 1753-1768.
18. Iglesias-Gato D, Wikström P, Tyanova S, Lavallee C, Thysell E, Carlsson J *et al.* The Proteome of Primary Prostate Cancer. *Eur Urol* 2016; **69**(5): 942-952.
19. Li QK, Chen J, Hu Y, Höti N, Lih T-SM, Thomas SN *et al.* Proteomic characterization of primary and metastatic prostate cancer reveals reduced proteinase activity in aggressive tumors. *Scientific Reports* 2021; **11**(1): 18936.
20. Mantsiou A, Makridakis M, Fasoulakis K, Katafigiotis I, Constantinides CA, Zoidakis J *et al.* Proteomics Analysis of Formalin Fixed Paraffin Embedded Tissues in the Investigation of Prostate Cancer. *Journal of Proteome Research* 2020; **19**(7): 2631-2642.
21. Kwon OK, Ha YS, Na AY, Chun SY, Kwon TG, Lee JN *et al.* Identification of Novel Prognosis and Prediction Markers in Advanced Prostate Cancer Tissues Based on Quantitative Proteomics. *Cancer Genomics Proteomics* 2020; **17**(2): 195-208.
22. Kmeťová Sivoňová M, Tatarková Z, Jurečková J, Kliment J, Híveš M, Lichardusová L *et al.* Differential profiling of prostate tumors versus benign prostatic tissues by using a 2DE-MALDI-TOF-based proteomic approach. *Neoplasma* 2021; **68**(1): 154-164.
23. Liss MA, Leach RJ, Sanda MG, Semmes OJ. Prostate Cancer Biomarker Development: National Cancer Institute's Early Detection Research Network Prostate Cancer Collaborative Group Review. *Cancer Epidemiology Biomarkers & Prevention* 2020; **29**(12): 2454.
24. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol* 2016; **40**(2): 244-252.
25. Wright JL, Salinas CA, Lin DW, Kolb S, Koopmeiners J, Feng Z *et al.* Prostate cancer specific mortality and Gleason 7 disease differences in prostate cancer outcomes between cases with Gleason 4 + 3 and Gleason 3 + 4 tumors in a population based cohort. *J Urol* 2009; **182**(6): 2702-2707.
26. Gearman DJ, Morlacco A, Chevillat JC, Rangel LJ, Karnes RJ. Comparison of Pathological and Oncologic Outcomes of Favorable Risk Gleason Score 3 + 4 and Low Risk Gleason Score 6 Prostate Cancer: Considerations for Active Surveillance. *J Urol* 2018; **199**(5): 1188-1195.
27. Dall'Era MA, Klotz L. Active surveillance for intermediate-risk prostate cancer. *Prostate Cancer Prostatic Dis* 2017; **20**(1): 1-6.
28. Hanna B, Ranasinghe W, Lawrentschuk N. Risk stratification and avoiding overtreatment in localized prostate cancer. *Curr Opin Urol* 2019; **29**(6): 612-619.
29. Alchin DR, Murphy D, Lawrentschuk N. Risk factors for Gleason Score upgrading following radical prostatectomy. *Minerva Urol Nefrol* 2017; **69**(5): 459-465.

30. Edge SB, Amin MB, Edge S, Greene F, Byrd DR, Brookland RK *et al.* *AJCC cancer staging manual 8th edition*, vol. 8. Springer New York, 2017.
31. Fair WR, Israeli RS, Heston WD. Prostate-specific membrane antigen. *Prostate* 1997; **32**(2): 140-148.
32. Barinka C, Hlouchova K, Rovenska M, Majer P, Dauter M, Hin N *et al.* Structural basis of interactions between human glutamate carboxypeptidase II and its substrate analogs. *J Mol Biol* 2008; **376**(5): 1438-1450.
33. Murphy GP, Elgamal AA, Su SL, Bostwick DG, Holmes EH. Current evaluation of the tissue localization and diagnostic utility of prostate specific membrane antigen. *Cancer* 1998; **83**(11): 2259-2269.
34. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clin Cancer Res* 1997; **3**(1): 81-85.
35. Frerker N, Wagner L, Wolf R, Heiser U, Hoffmann T, Rahfeld JU *et al.* Neuropeptide Y (NPY) cleaving enzymes: structural and functional homologues of dipeptidyl peptidase 4. *Peptides* 2007; **28**(2): 257-268.
36. Wesley UV, McGroarty M, Homoyouni A. Dipeptidyl Peptidase Inhibits Malignant Phenotype of Prostate Cancer Cells by Blocking Basic Fibroblast Growth Factor Signaling Pathway. *Cancer Research* 2005; **65**(4): 1325.
37. Pro B, Dang NH. CD26/dipeptidyl peptidase IV and its role in cancer. *Histol Histopathol* 2004; **19**(4): 1345-1351.
38. Ueda K, Tatsuguchi A, Saichi N, Toyama A, Tamura K, Furihata M *et al.* Plasma low-molecular-weight proteome profiling identified neuropeptide-Y as a prostate cancer biomarker polypeptide. *J Proteome Res* 2013; **12**(10): 4497-4506.
39. Compagno D, Gentilini LD, Jaworski FM, Pérez IG, Contrufo G, Laderach DJ. Glycans and galectins in prostate cancer biology, angiogenesis and metastasis. *Glycobiology* 2014; **24**(10): 899-906.
40. Merseburger AS, Kramer MW, Hennenlotter J, Simon P, Knapp J, Hartmann JT *et al.* Involvement of decreased Galectin-3 expression in the pathogenesis and progression of prostate cancer. *Prostate* 2008; **68**(1): 72-77.
41. Knapp JS, Lokeshwar SD, Vogel U, Hennenlotter J, Schwentner C, Kramer MW *et al.* Galectin-3 expression in prostate cancer and benign prostate tissues: correlation with biochemical recurrence. *World J Urol* 2013; **31**(2): 351-358.
42. Ahmed H, Banerjee PP, Vasta GR. Differential expression of galectins in normal, benign and malignant prostate epithelial cells: silencing of galectin-3 expression in prostate cancer by its promoter methylation. *Biochem Biophys Res Commun* 2007; **358**(1): 241-246.
43. Capuani F, Conte A, Argenzio E, Marchetti L, Priami C, Polo S *et al.* Quantitative analysis reveals how EGFR activation and downregulation are coupled in normal but not in cancer cells. *Nature Communications* 2015; **6**(1): 7999.
44. Barinka C, Nagappa AN, Bhatt S, Kanoujia J. Studies on Structures and Functions of Kinases leading to Prostate Cancer and Their Inhibitors. *Current Enzyme Inhibition* 2020; **16**(1): 90.

45. Nastały P, Stoupiec S, Popęda M, Smentoch J, Schlomm T, Morrissey C *et al.* EGFR as a stable marker of prostate cancer dissemination to bones. *British Journal of Cancer* 2020; **123**(12): 1767-1774.
46. Rossini A, Giussani M, Ripamonti F, Aiello P, Regondi V, Balsari A *et al.* Combined targeting of EGFR and HER2 against prostate cancer stem cells. *Cancer Biol Ther* 2020; **21**(5): 463-475.

## Tables

**Table 1.** Summary of primary antibodies

#	Antibodies	Resource	Clone	Clonality	Dilution	Catalog #
1	DPP4/CD26	Cell Signaling	D6D8K	Rabbit Monoclonal	1:200	67138T
2	Total EGFR	Cell Signaling	D38B1	Rabbit monoclonal	1:50	4267
3	Phospho-EGFR (phospho Y1068)	Cell Signaling	D7A5	Rabbit Monoclonal	1:1,000	3777
4	MAN1B1	Sigma-Alderich	1D6	Mouse Monoclonal	1:100	MABS1222
5	PSMA	Dako	3E6	Mouse Monoclonal	1:100	M3620
6	P16	Ventana	INK4a	Mouse Monoclonal	Prediluted by manufacturer	705-4793
7	Galectin-3	Ventana	9C4	Mouse Monoclonal	Prediluted by manufacturer	790-4256
8	PD-1	Cell Marque	NAT105	Mouse Monoclonal	1:100	315M-96
9	PD-L1 (SP142)	Spring Bioscience	SP142	Rabbit Monoclonal	1:100	M4422
10	PD-L1 (22C3)	Dako	22C3	Mouse Monoclonal	1:100	SK006
11	AR	Cell Marque	SP107	Rabbit Monoclonal	Prediluted by manufacturer	200R-18
12	PTEN	Biocare Medical	6H2.1	Mouse Monoclonal	1:100	PM278AA

N/A: not applicable. PSMA: prostate specific membrane antigen. AR: androgen receptor.

**Table 2.** Clinical features of prostate cancer patients

Characteristics	Prostate adenocarcinoma
<b>Age(years)</b>	
Median	61
Range	40-73
<b>Tumor Location (n=57)</b>	
Left side	20 (35.09%)
Right side	36 (63.16%)
N/A	1 (1.75%)
<b>Pathological stage (n=57)</b>	
pT2	30 (52.63)
pT2X	1 (1.75%)
pT3A	13 (22.81%)
pT3B	11 (19.30%)
pT4	2 (3.51%)

N/A: not applicable.

**Table 3.** IHC staining scores of protein markers

Markers	IHC Staining Scores	NAT	Gleason 3	Gleason 4	Gleason 5
AR	Tissue Cores (n)	117	92	61	57
	Mean ± SD	2.21 ± 0.83	2.53 ± 0.69	2.61 ± 0.71	2.75 ± 0.71
DPP4	Tissue Cores (n)	113	92	63	62
	Mean ± SD	2.51 ± 0.75	2.66 ± 0.62	1.92 ± 0.94	1.40 ± 0.91
Galectin-3	Tissue Cores (n)	117	91	62	62
	Mean ± SD	2.16 ± 0.84	1.43 ± 0.80	1.16 ± 0.93	0.85 ± 0.88
MAN1B1	Tissue Cores (n)	121	92	61	63
	Mean ± SD	1.90 ± 0.97	2.40 ± 0.66	2.30 ± 0.78	2.00 ± 0.92
P16	Tissue Cores (n)	112	92	62	61
	Mean ± SD	0.33 ± 0.54	1.10 ± 0.89	1.34 ± 0.83	1.08 ± 0.82
Phospho-EGFR	Tissue Cores (n)	N/A	92	63	63
	Mean ± SD	N/A	1.58 ± 0.63	1.83 ± 0.79	1.98 ± 0.87
PSMA	Tissue Cores (n)	126	92	63	60
	Mean ± SD	0.96 ± 0.87	1.54 ± 0.79	2.40 ± 0.91	2.48 ± 0.95

N/A: not applicable. PSMA: prostate specific membrane antigen. AR: androgen receptor.

## Figures

Figure 1

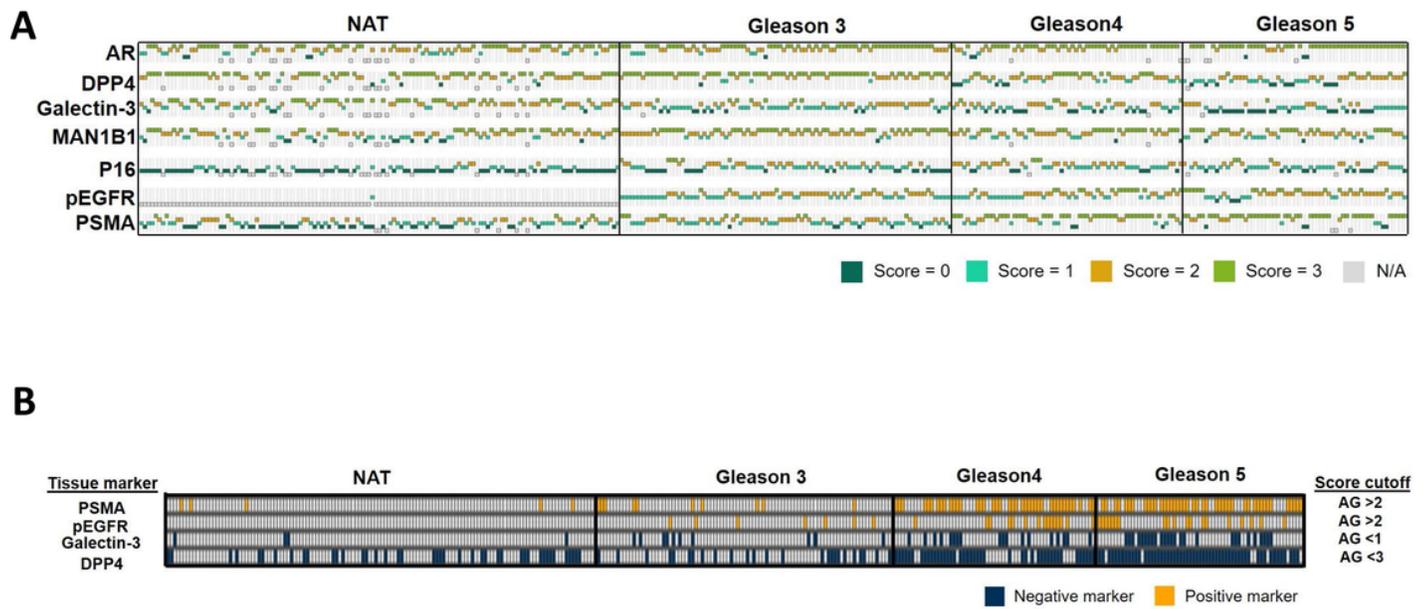
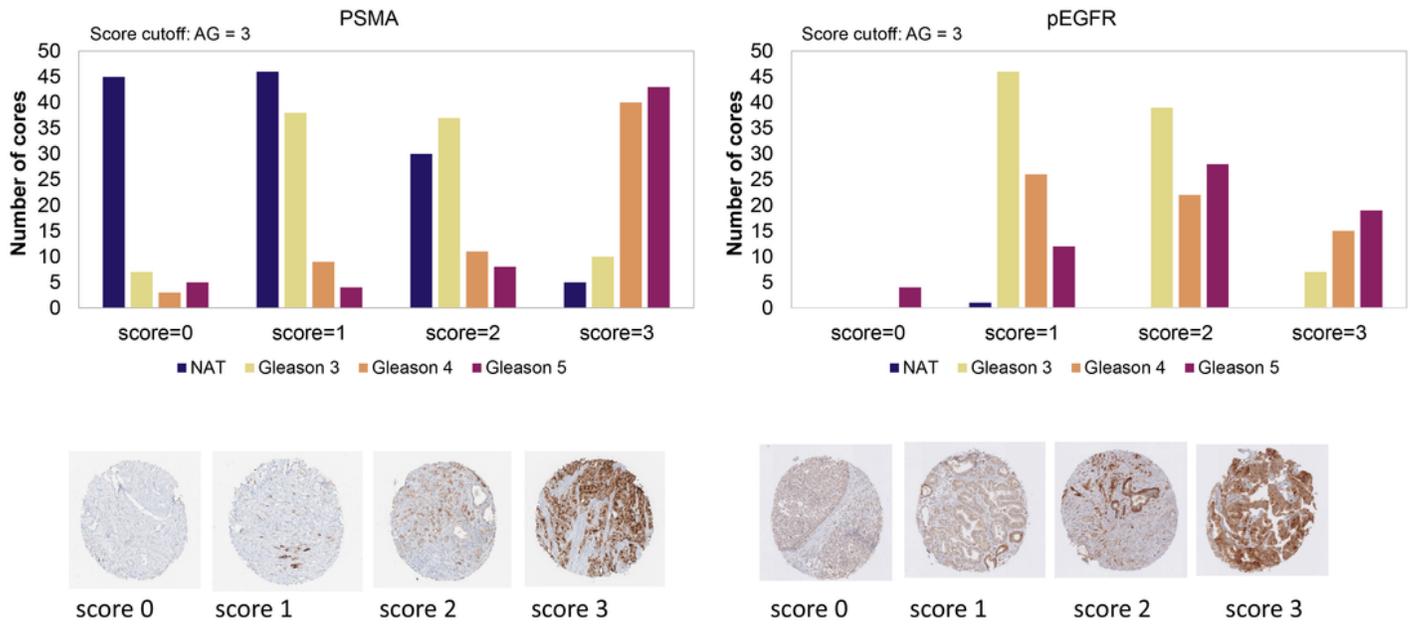


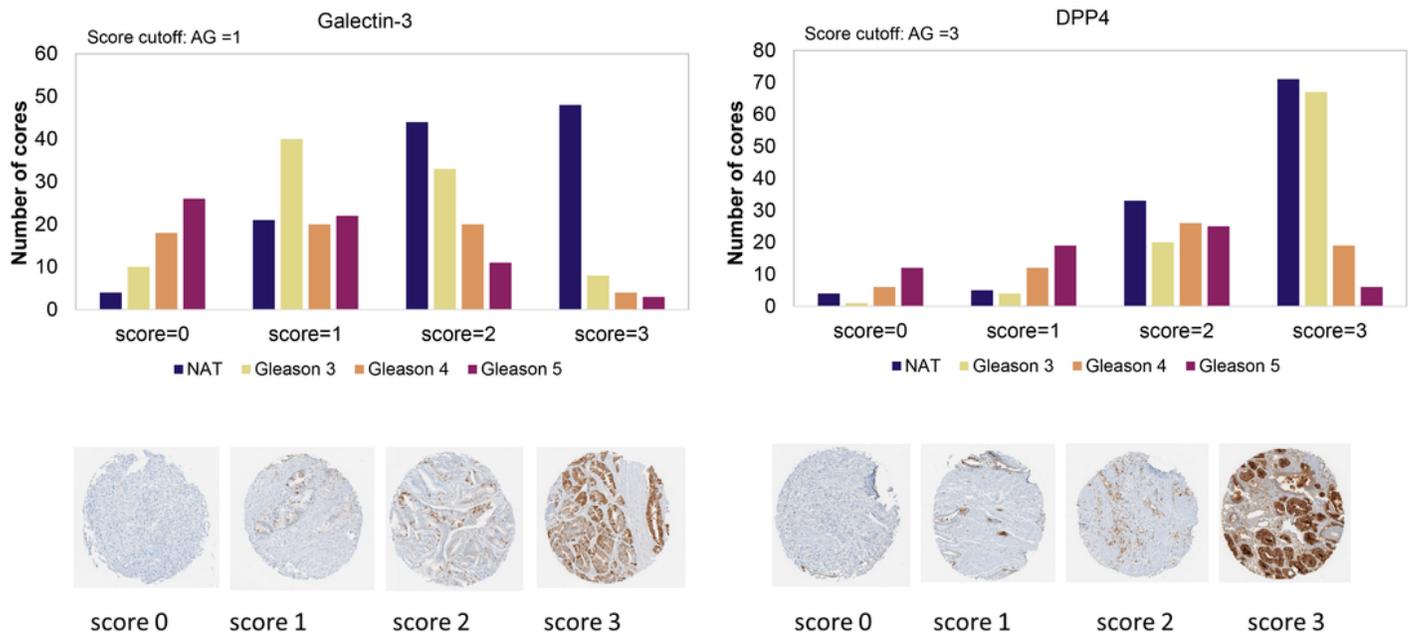
Figure 1

Heatmap of staining patterns of individual protein in PCa TMA. A. overall staining patterns. B. Individual marker in the separation of AG from NAG tumors using different cut-off Gleason scores. The intensity of IHC stains were scored using a semi-quantitative 4-tier scoring system: 0 (0%, no staining), 1 (<10%, weak and focally staining), 2 (10-50%, medium and focally staining), or 3 (>50%, strong and diffusely staining) in tumor cells. AR: androgen receptor. pEGFR: phospho-EGFR. PSMA: prostate specific membrane antigen.

**Figure 2A**



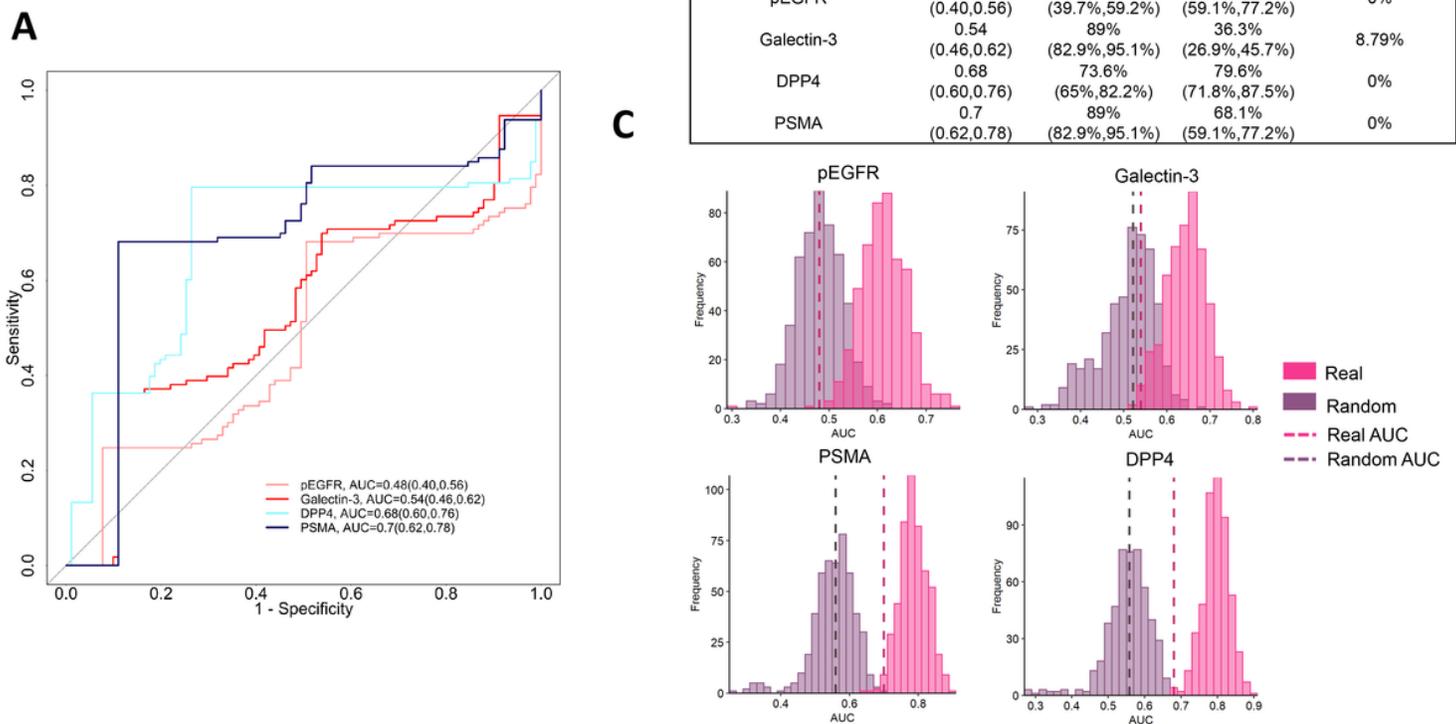
**Figure 2B**



**Figure 2**

IHC staining patterns of individual protein markers. A. overall staining patterns of PSMA and phosphor-EGFR (pEGFR). B. overall staining patterns of Galectin-3 and PDD4. The intensity of IHC stains were scored using a semi-quantitative 4-tier scoring system: 0 (0%, no staining), 1 (<10%, weak and focally staining), 2 (10-50%, medium and focally staining), or 3 (>50%, strong and diffusely staining) in tumor cells.

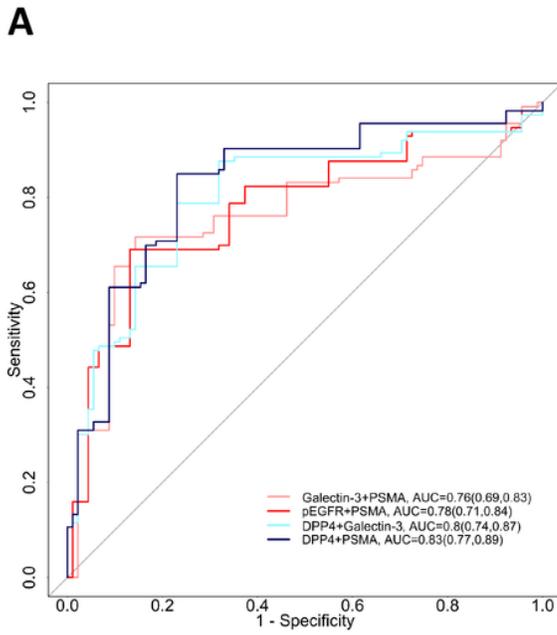
**Figure 3**



**Figure 3**

Receiver operating characteristic (ROC) analysis of individual protein marker. A. Comparison of individual AUC curve. B. Specificities and sensitivities of individual protein marker. C. Comparison of values between real and random (label permutation analysis) of individual protein marker.

Figure 4



**B**

Tissue marker panel	AUC (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity at 95% sensitivity
Galectin-3+PSMA	0.76 (0.69,0.83)	85.7% (78.9%,92.5%)	71.7% (62.9%,80.5%)	7.69%
pEGFR+PSMA	0.78 (0.71,0.84)	86.8% (80.2%,93.4%)	69% (60%,78%)	6.59%
DPP4+Galectin-3	0.8 (0.74,0.87)	68.1% (59%,77.2%)	87.6% (81.2%,94%)	4.40%
DPP4+PSMA	0.83 (0.77,0.89)	76.9% (68.7%,85.1%)	85% (78%,91.9%)	38.46%

**C**

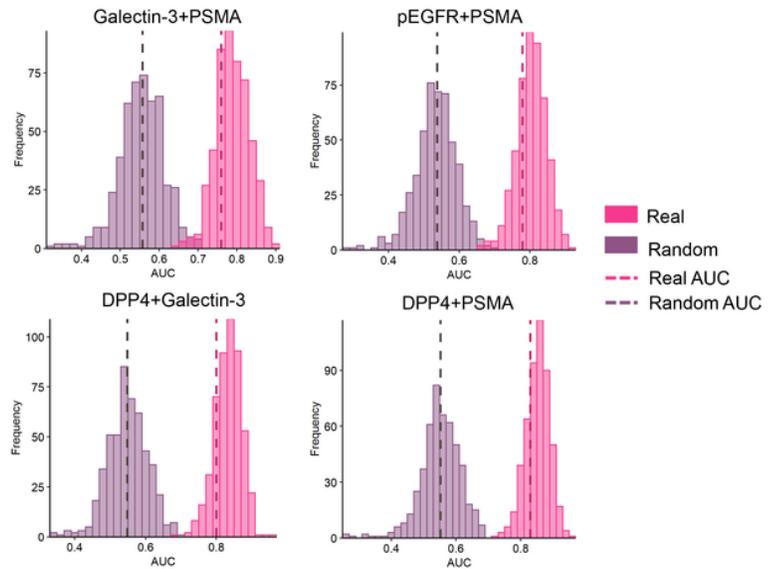
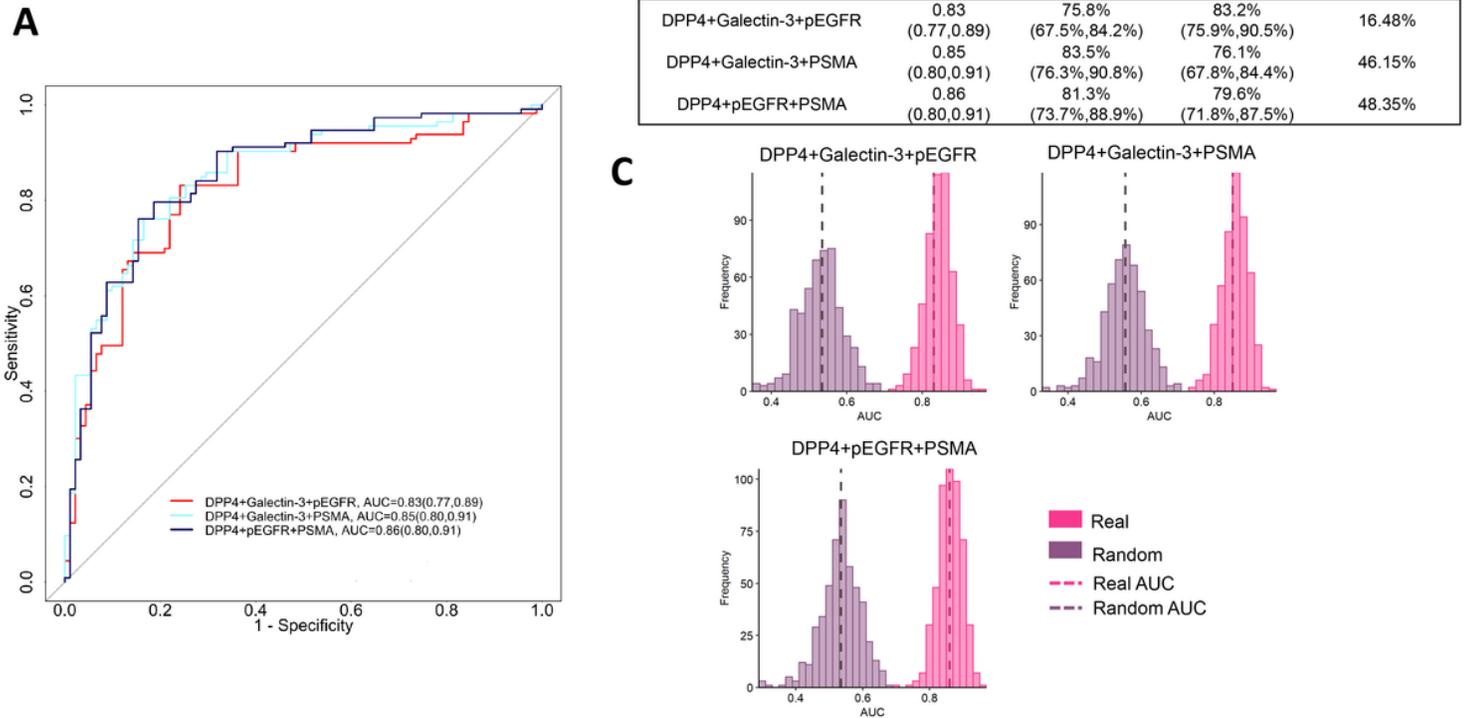


Figure 4

Receiver operating characteristic (ROC) analysis of two-marker panels, containing Galectin-3 plus PSMA, phospho-EGFR plus PSMA, DPP4 plus PSMA and Galectin-3 plus DPP4. A. Comparison of AUC curves of different panels. B. Specificities and sensitivities of two-marker panels. C. Comparison of values between real and random (label permutation analysis) of different panels.

**Figure 5**



**Figure 5**

Receiver operating characteristic (ROC) analysis of three-marker panels, containing Galectin-3 plus phospho-EGFR plus DPP4, Galectin-3 plus PSMA plus DPP4, and phospho-EGFR plus PSMA plus DPP4. A. Comparison of AUC curves of different panels. B. Specificities and sensitivities of three-marker panels. C. Comparison of values between real and random (label permutation analysis) of different panels.

## Supplementary Files

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

- [1132021Supplementarytableclinicalinfo.xlsx](#)