ORIGINAL ARTICLE

The Prostate WILEY

Molecular classification of hormone-sensitive and castration-resistant prostate cancer, using nonnegative matrix factorization molecular subtyping of primary and metastatic specimens

Kobe C. Yuen PhD[1] | Ben Tran MBBS[2,3] | Angelyn Anton MBBS[3,4] |
Habib Hamidi PhD[1] | Anthony J. Costello MBBS, MD[5,6,7] |
Niall M. Corcoran MB, BCh, BAO, PhD[5,6] | Nathan Lawrentschuk MBBS, PhD[3,5,6] |
Natalie Rainey MBBS[3] | Marie C. G. Semira MBBS[3] | Peter Gibbs MBBS, MD[3] |
Sanjeev Mariathasan PhD[1] | Shahneen Sandhu MBBS[2] | Edward E. Kadel III BSc[1]

[1]Department of Oncology Biomarker Development, Genentech, Inc., South San Francisco, California, USA

[2]Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Victoria, Australia

[3]Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia

[4]Eastern Health, Melbourne, Victoria, Australia

[5]Royal Melbourne Hospital, Melbourne, Victoria, Australia

[6]Department of Surgery, The University of Melbourne, Melbourne, Victoria, Australia

[7]Australian Prostate Centre, North Melbourne, Victoria, Australia

Correspondence
Edward E. Kadel III, US Medical Affairs & OBD, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Email: ekadel@gene.com

Abstract

Background: Despite the rapidly evolving therapeutic landscape, immunotherapy has demonstrated limited activity in prostate cancer. A greater understanding of the molecular landscape, particularly the expression of immune-related pathways, will inform future immunotherapeutic strategies. Consensus nonnegative matrix factorization (cNMF) is a novel model of molecular classification analyzing gene expression data, focusing on biological interpretation of metagenes and selecting meaningful clusters.

Objective: We aimed to identify molecular subtypes of prostate cancer using cNMF and correlate these with existing biomarkers to inform future immunotherapeutic strategies.

Methods: A cohort of archival tumor specimens from hormone-sensitive and castration-resistant disease was studied. Whole transcriptomic profiles were generated using TruSeq RNA Access technology and subjected to cNMF. Comprehensive genomic profiling was performed with the FoundationOne assay. NMF subtypes were characterized by gene expression pathways, genomic alterations and correlated with clinical data, then applied to The Cancer Genome Atlas data set.

Results: We studied 164 specimens, including 52 castration-resistant and 13 paired primary/metastatic specimens. cNMF identified four distinct subtypes. NMF1 (19%) is enriched for immune-related and stromal-related pathways with transforming growth factor β (TGFβ) signature. NMF2 (36%) is associated with FOXO-mediated transcription signature and AKT signaling, NMF3 (26%) is enriched for ribosomal

---

Kobe C. Yuen, Ben Tran, Angelyn Anton, Habib Hamidi are co-first authors for this study.

Shahneen Sandhu and Edward E. Kadel are co-last authors for this study.

RNA processing, while NMF4 (19%) is enriched for cell cycle and DNA-repair pathways. The most common gene alterations included TMPRSS22 (42%), TP53 (23%), and DNA-repair genes (19%), occurring across all subtypes. NMF4 is significantly enriched for MYC and Wnt-signaling gene alterations. TMB, CD8 density, and PD-L1 expression were low overall. NMF1 and NMF4 were NMF2 was associated with superior overall survival.

Conclusions: Using cNMF, we identified four molecularly distinct subtypes which may inform treatment selection. NMF1 demonstrates the most inflammatory signature with asuppressive TGFβ signature, suggesting potential benefit with immunotherapy combination strategies targeting TGFβ and PD-(L)1. Prospective studies are required to evaluate the use of this novel model to molecularly stratify patients for optimal treatment selection.

KEYWORDS
CRPC, HSPC, molecular subtyping, TME

1 | INTRODUCTION

The therapeutic landscape in advanced prostate cancer is rapidly evolving following development of multiple life-prolonging treatments.[1] In contrast, immune checkpoint inhibition has demonstrated limited activity in prostate cancer, despite unprecedented survival benefits in other malignancies. The biological underpinning for this limited antitumour activity remains poorly understood and islikely multifactorial including an immunosuppressive tumor microenvironment, low PD-L1 expression, and low tumor mutation burden (TMB).[2,3] Previous studies suggest activity in certain subsets including those with high PD-L1 expression and high TMB or specific genomic alterations involving mismatch or DNA repair genes.[4–7] Greater understanding of the molecular landscape in prostate cancer, namely the expression of immune-related pathways, may inform patient selection and lead to new immunotherapeutic strategies.

Several methods have historically been utilized for genomic or molecular subtyping of prostate cancer. Gene expression-based subtyping using the PAM50 model developed for breast cancer[8] was subsequently applied to prostate cancer, distinguishing luminal and basal subtypes.[9] However, a de novo prostate cancer-specific model that features immune pathways has not been reported.

Consensus nonnegative matrix factorization (cNMF) is an unsupervised machine learning approach for identifying transcriptomic-based molecular subtypes using gene expression.[10,11] cNMF has been utilized in several tumor types, including prostate cancer, using data from The Cancer Genome Atlas (TCGA), demonstrating feasibility in grouping genetic information and identifying metagenes correlating with tumor subtypes or characteristic genomic alterations.[12,13]

Our study aimed to evaluate distinct molecular subtypes of prostate cancer using cNMF clustering in a cohort of primary and metastatic tumor specimens spanning hormone-sensitive (HSPC) and castration-resistant (CRPC) states. Established biomarkers including TMB, PD-L1 expression, CD8 T-cell density, genomic alterations, and immune-related signatures were overlaid onto NMF subtypes and correlated with clinicopathological features, patient demographics, and clinical outcomes.

2 | MATERIALS AND METHODS

2.1 | Patient cohort

A selected cohort of 164 formalin-fixed paraffin-embedded (FFPE) archival prostate cancer specimens was analyzed in this study including primary prostate and metastatic specimens across both HSPC and CRPC disease states (Figure 1 and Table 1). Tumor specimens were identified using institutional pathology and prostate cancer databases from two institutions and had been acquired between 2000 and 2015. Retrospective review of medical records provided clinicopathological details including disease state (HSPC/CRPC), age, Gleason score, prior therapies, clinical follow-up, and overall survival data. Institutional ethical approval was granted for the analysis of archival tissue and clinical data from this cohort.

FFPE tissue was macro-dissected for the tumor area. RNA was extracted using the High Pure FFPET RNA Isolation Kit (Roche) and assessed by Qubit and Agilent Bioanalyzer for quantity and quality. Whole transcriptome profiles were generated using TruSeq RNA Access technology. Specimens without sufficient tumor content were excluded from RNAseq analysis.

2.2 | cNMF

We sought to identify transcriptomic subtypes by applying cNMF to gene expression data from $n = 94$ primary tumor specimens only
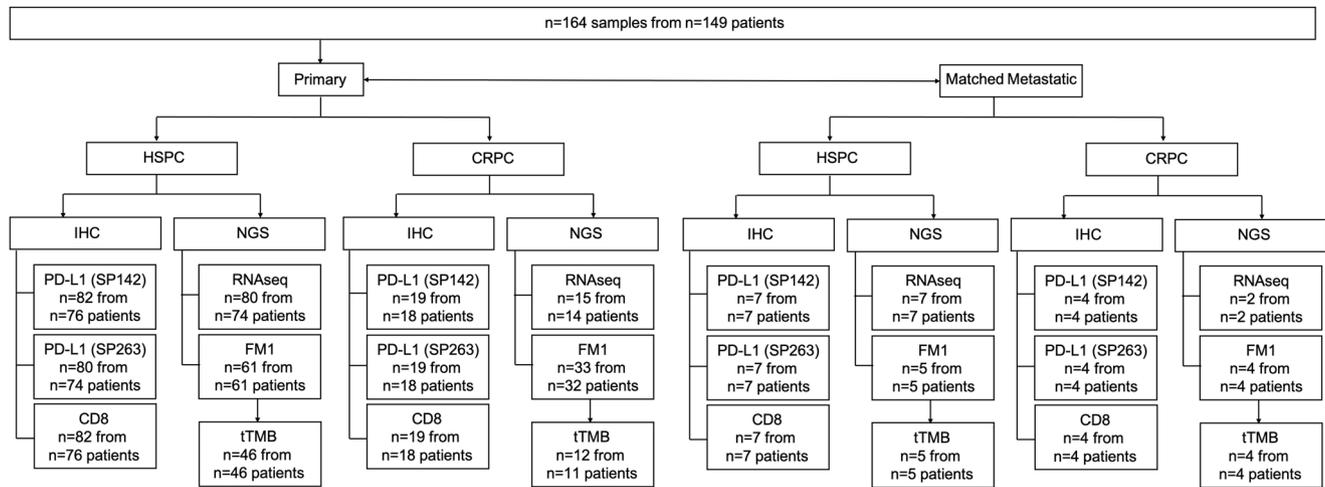
FIGURE 1 Samples for biomarker analysis. Flowchart showing the number of primary and matched metastatic samples with immunohistochemistry (IHC) and next-generation sequencing (NGS) included for analysis. CRPC, castrate-resistant prostate cancer; HSPC, hormone-sensitive prostate cancer

(n = 80 HSPC and n = 14 CRPC), given primary tissue is most commonly available in men with metastatic prostate cancer. We selected 2804 genes (top 10%) with the highest variability across patients, using median absolute deviation analysis (CRAN. R package version 0.22.0).[11] cNMF computes multiple k-factor factorization decompositions of the expression matrix and evaluates the stability of the solutions using a cophenetic coefficient. We tested k2–8 clusters and used the maximal optimization of the cophenetic score to determine the number of clusters (Figure S1).

To further validate our molecular subtyping schema, we used the random forest machine learning algorithm (R package) to derive a classifier and predict cNMF clusters in an independent data set (TCGA) including 333 primary prostate specimens, and inthe metastatic specimens from our cohort.[14] We limited the gene expression matrix in the test and training set to the top 10% of most variable genes used in the NMF clustering (n = 3072) above.

To show the advantages of cNMF over existing subtyping methods, we also classified the primary tumors from our cohort using PAM50 subtyping as previously described[8,9] and compared biological features of interest in this study.

2.3 | Differential gene expression by gene set enrichment analysis (GSEA)

GSEA was performed by Camera enrichment method in the multiGSEA R package to compare NMF subtypes, with the use of the Reactome gene set collections from the Molecular Signature Database.[15] Pathway Z scores were calculated for each gene set using the scores function. Deconvolution analyses for tumor samples were performed using the xCell cell types enrichment score tool (http://xcell.ucsf.edu/)[16] and the MCPcounter R package.[17]

2.4 | Genomic profiling

Comprehensive next-generation sequencing (NGS) was carried out using the FoundationOne (FM1) assay (T7 baitset; Foundation Medicine Inc.) that profiles 395 genes. We assessed all classes of genomic alterations including short variants, deletions, amplifications, and rearrangements, with only known (COSMIC[18]) or likely alterations considered for analysis, as described previously.[19]

2.5 | Immunohistochemistry: PD-L1 and CD8

FFPE tumor tissue (4 uM sections) was stained for PD-L1 by immunohistochemistry (IHC) using Roche Diagnostics (Ventana) anti-human PD-L1 monoclonal antibodies (SP142 or SP263). Samples were scored for PD-L1 expression on tumor cells and surrounding immune-infiltrating cells as IC0, 1, 2, or 3 if <1%, ≥1% but <5%, ≥5% but <10%, or ≥10% of cells were PD-L1 positive, respectively.

Tumors were stained for CD8 IHC with a mouse monoclonal antibody (clone C8/144B) on 4 uM FFPE sections. Whole slide images of the CD8-stained tumors were then analyzed with Definiens, where the tumor regions of interest were generated. The area fraction (%marker area) of the CD8-stained area was then digitally calculated.

2.6 | Statistical analysis

All analyses were conducted using the R package (version 3.6.1; R Foundation for Statistical Computing). Unless otherwise stated, all comparisons for continuous variables use two-sided Mann–Whitney U tests for two groups and Kruskal–Wallis tests for over two groups. p Values are reported as *p < 0.05; **p < 0.01; ***p < 0.001. Survival

TABLE 1 Clinicopathological features

	All	HSPC	CRPC	Unknown
Number of specimens	164	106	52	6
Number of patients	149[a]	98[a]	51[a]	6
Median age	66.0	64.4	69.7	75.9
Specimen sites				
Prostate	103	83	20	0
Bone	24	6	13	5
Pelvic lymph node	9	9	0	0
Distant lymph node	8	6	2	0
Liver	3	0	3	0
Lung	4	2	2	0
Soft tissue	10	0	10	0
Brain	3	0	2	1
Other	0	0	0	0
Gleason score[b]				
6	3	3	0	0
7	46	46	0	0
8	10	10	0	0
9	38	38	0	0
10	5	5	0	0
Unknown	8	2	0	6
Not available	54	2	52	0
Treatment received before specimen collection				
Androgen deprivation therapy	53	3	50	0
Chemotherapy	6	0	6	0
Abiraterone/enzalutamide	1	0	1	0
Year of specimen				
2000–2004	12	4	7	1
2005–2009	92	67	21	4
2010–2015	60	35	24	1

[a]There were eight patients with two separate Hormone-Sensitive Prostate Cancer (HSPC) specimens, one patient with two separate castrate-resistant prostate cancer (CRPC) specimens, six patients with paired HSPC/CRPC specimens, and 12 patients with paired primary/metastatic specimens.

[b]Gleason grade cannot be determined in specimens where ADT had already been started.

analyses were conducted using Cox-proportional hazard models. Overall survival was calculated from the date of prostate cancer diagnosis to the date of death. Alive patients were censored at the date of the last follow-up. We constructed Kaplan–Meier curves to visualize overall survival, stratified by NMF subtypes, and compared groups using log-rank tests.

3 | RESULTS

A total of 164 tumor specimens were retrieved from 149 patients: 103 primary prostate and 61 metastatic specimens, including 52 CRPC and 106 HSPC (Table 1). The majority of metastatic sites were bone ($N = 24$) or lymph node ($N = 17$) sites. There were 13 patients with paired primary/metastatic specimens, including seven HSPC and six CRPC metastatic specimens. Median follow-up from the date of prostate cancer diagnosis was 12.2 years. Table 1 details Gleason scores and prior systemic treatments. As shown in Figure 1, RNAseq was successful in 94 primary specimens (91%), NGS in 111 specimens (68%), CD8 IHC in 138 specimens (84%), and PD-L1 IHC (SP263) in 158 specimens (96%).

cNMF was applied to transcriptomes of 94 primary tumor samples ($n = 80$ HSPC and $n = 14$ CRPC), the most robust clustering was identified for $k = 4$, leading to identification of four molecular subtypes of prostate cancer, NMF1 ($n = 18$, 19%), NMF2 ($n = 34$, 36%), NMF3 ($n = 24$, 26%) and NMF4 ($n = 18$, 19%) (Figure S1). RNA quality was consistent across all subtypes. There was no significant difference in total reads and total analyzed reads between the subtypes (data not shown). NMF1 is enriched for immune-related pathways such as PD-L1 and activated T-cell signaling pathways including interleukin-2 and interferon-γ (Figure 2A,B). Additionally, NMF1 has strong enrichment for stromal-related pathways. Using unbiased GSEA analysis, NMF1 had significantly higher immune signatures involving T-effector, interferon-γ, immune checkpoint, and macrophages (Figures 2C,D and S2). Deconvolution analyses demonstrate higher infiltration of immune cells including monocytes, CD8, and dendritic cells in NMF1 (Figure S3A). The transforming growth factor β (TGFβ) signature is also significantly enriched in NMF1, characterized by the enrichment for fibroblasts and smooth muscle cells (Figure 3A,B).

NMF2 subtype is associated with FOXO-mediated transcription signature and regulation of PI3Kinase/AKT signaling. NMF3 subtype is enriched for ribosomal RNA processing and translation, whereas NMF4 subtype is enriched for cell cycle, DNA repair pathways, and androgen receptor signature (Figures 2C–E and S2). NMF3 and four demonstrate lower expression of PI3Kinase/AKT-associated genes (Figures 2C–E and S2). NMF4 is enriched for CRPC samples, while NMF3 and NMF4 were enriched for metastatic specimens (Figure 2C). NMF2 were associated with superior overall survival (OS) compared to other NMF subtypes ($p = 2.13e-4$) (Figure 2F).

PD-L1 expression was low within most primary tumors; 65 (69.1%) tumors were IC0, while 3 (3.2%) were IC2 by SP142 assay and there were no significant differences in PD-L1 IC scores between cNMF subtypes (Figure S4A,B). NMF3 and NMF4 subgroups had higher Gleason score compared to NMF1 and NMF2 ($p = 1.7e-7$) (Figure S4C); however, CD8 density was not significantly different between cNMF subtypes (Figure S4D).

When cNMF was applied to metastatic specimens in our cohort, all were classified within NMF3 or NMF4 subtypes (Figure 3A). Notably, 11 of the 13 cases with paired primary and metastatic
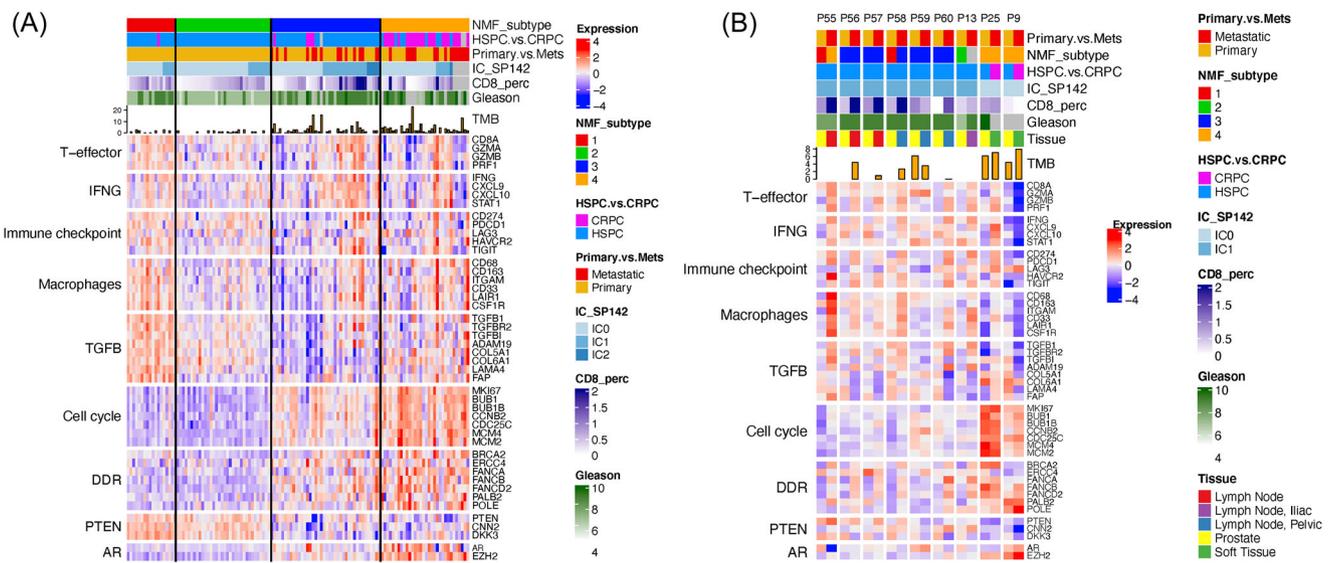
FIGURE 2 Distinct molecular signatures in each NMF subtype that is associated with overall survival. (A) Unbiased Reactome pathway enrichment analysis depicts distinct biology associated with each NMF subtype. (B) Heatmap shows mean expression levels of some enriched pathways in (A). (C) Heatmap shows representative genes from different gene signatures. Red, high expression; blue, low expression. (D) Boxplot shows the expression levels of the gene signatures in (C). Statistical significance between NMF subtypes is calculated by Mann–Whitney U test, $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$. (E) Radar plot shows the overall patterns of gene signatures in (C). (F) Kaplan–Meier curves showing patients with available diagnosis follow-up data and overall survival stratified by NMF subtypes. p Value is computed by log-rank test. NMF, nonnegative matrix factorization; TMB, tumor mutation burden [Color figure can be viewed at wileyonlinelibrary.com]

specimens had consistent NMF classification. Only two cases had discordance: one had NMF1 classified primary and NMF4 classified metastatic specimen, while another patient had NMF4 classified primary and NMF3 classified metastatic specimen (Figures 3B and S5A). The metastatic specimens had numerically higher IC scores compared to primary specimens ($p = 0.45$): IC1 (8/24 = 35% vs. 26/94 = 28%) and IC2 (3/24 = 13% vs. 3/94 = 3%) (Figure S5B). The majority of CRPC specimens were IC0 including primary (12/14 = 86%) and metastatic (8/9 = 89%) sites. Most primary

specimens had low CD8 cell density (median: 0.48%), with 116 (84.7%) being ≤1%. Although CD8 density was not significantly different between cNMF subtypes (Figure S2D), CD8 density was higher in metastatic specimens (median 1.31), compared to primary HSPC samples (Median 0.51) (Figure S5C). However, most metastatic specimens were lymph nodes ($N = 17$), where CD8 density is expected to be high as a result of the surrounding nodal tissue. Notably, primary ($p = 0.039$) and metastatic ($p = 0.014$) HSPC specimens had significantly higher CD8 densities compared to CRPC specimens.

FIGURE 3 Biomarkers between hormone-sensitive prostate cancer (HSPC) and matched castrate-resistance prostate cancer (CRPC). (A) Heatmap shows expression of gene signatures in Figure 2C with the addition of metastatic samples. (B) Heatmap shows gene signatures with matched HSPC and CRPC samples from nine patients with RNAseq data. [Color figure can be viewed at wileyonlinelibrary.com]

Our cNMF clusters demonstrated concordant gene expression patterns for each NMF subtype in both our cohort and that of the TCGA (Figure 4A,B). Within the TCGA cohort, NMF1 was characterized by high expression of immune genes, TGFβ/stromal-related genes, and PI3Kinase/AKT-activated signature. The NMF4 subgroup demonstrated the highest expression of the cell cycle, DDR pathway, and AR-related genes (Figure 4C,D).

We also performed PAM50 subtyping[8] (Figure S6A) and compared it with the cNMF subtyping. While immune signatures such as T-effector, interferon-γ and macrophages, as well as TGFβ signature, showed significant enrichment in the cNMF subgroups (Figure 2D), there was no significant enrichment of these signatures across the three PAM50 subtypes (Figure S6B,C). However, LumB subtype was enriched for the cell cycle, DDR, and AR signatures but not the PTEN signature (Figure S6B,C). Basal and LumB subtypes were associated with the best and worst OS, respectively, in the PAM50 subtyping (Figure S6D).

The FM1 Assay identified transmembrane serine protease 2 (TMPRSS2) as the most commonly altered gene (42%) in this cohort (Figure S7). Alterations in TP53 (23%) and DNA repair-associated tumor suppressor genes (19%) such as BRCA2 and PTEN were also common (Figure S7).

Among the most commonly altered genes (≥5%), 20 of 62 (32.2%) occur in all four NMF subtypes (Figure S9A,B). Comprehensive genomic alteration profiles and somatic variants by NMF subtype are demonstrated in Figure 5A,B as well as Figure S8–S11, respectively. NMF4 is significantly enriched for oncogene MYC ($p = 0.037$) and Wnt-signaling genes such as APC ($p = 0.014$) and TSC2 ($p = 0.024$) (Figure S9C). Wnt ($p = 0.029$) and SHH ($p = 0.046$) signaling pathways were significantly enriched in NMF4 (Figure S12). SHH pathway alterations were primarily short variant alterations of Protein patched homolog 2 (PTCH2) (Figures 5B and S11A). No other

pathways or signatures were significantly different between NMF subtypes.

Median TMB in our cohort was low (1.75 mutations/Mb). NMF4 samples have the highest TMB (median: 2.63 mutations/Mb), while NMF2 have the lowest (median: 0.88 mutations/Mb) (Figure 5C). Tumor samples with alterations in the homologous recombination repair (HRR) pathway demonstrated significantly higher TMB compared to the overall cohort (Figure S12A,B), however within each NMF subtype there was no difference (Figures S13 and S14). GSEA demonstrated that genes related to the mitotic cell cycle, DNA repair, and translation were significantly associated with higher TMB (Figure 5D). TMB was not statistically different between CRPC and HSPC settings, nor between primary and metastatic specimens (Figure S15).

4 | DISCUSSION

In this study, we describe a novel classification system utilizing cNMF subtyping in a cohort of both HSPC and CRPC and compare it with established biomarkers. By including both disease states from patient samples with minimal treatment before collection, we sought to compare the inherent biology for both HSPC and CRPC within the same analyses. We molecularly stratified prostate cancer into four distinct subtypes. NMF1 demonstrated enrichment of stromal- and immune-related pathways, including PD-L1 and T-cell signaling. In contrast, the NMF4 subtype was significantly enriched for DNA-repair pathways, androgen receptor gene signature, and increased TMB. As expected, CRPC samples were over-represented within this subtype, supporting previous observations of enrichment for somatic alterations in the DNA repair and androgen receptor signaling pathways in CRPC.[20] Metastatic specimens were also enriched

FIGURE 4 Distinct molecular signatures in each NMF subtype in The Cancer Genome Atlas. (A) Unbiased Reactome pathway enrichment analysis depicts distinct biology associated with each NMF subtype. (B) Heatmap shows mean expression levels of some enriched pathway in (A). (C) Representative genes from different gene signatures. Red, high expression; blue, low expression. (D) Radar plot shows the overall patterns of gene signatures in (C). NMF, nonnegative matrix factorization; TMB, tumor mutation burden [Color figure can be viewed at wileyonlinelibrary.com]

within NMF4, consistent with previous literature identifying cell cycle control and DNA damage response as key pathways in progression to metastatic disease.[21] It is noteworthy that paired primary and metastatic specimens had mostly consistent NMF classification, lending support for the clinical utility of cNMF classification using FFPE tumor material from the original diagnosis.

We applied our model to the TCGA cohort and demonstrated consistency between subtypes. Although previous classification models have been utilized in prostate cancer, including unsupervised hierarchical clustering[22,23] and self-organizing map methods[24], cNMF represents a novel subtyping methodology with several advantages. It is an efficient method for the identification of distinct molecular patterns. Additionally, cNMF has a higher resolution than hierarchical clustering and is more stable than self-organizing maps.[11,25] When compared to PAM50, which is often used in prostate cancer despite being developed in breast cancer, our study has shown the advantages of cNMF, particularly toward the clustering of biological features related to cancer immunology. For example, significant

enrichment of immune and TGFβ signatures was in the NMF1 subtype while no such enrichment was observed across the three PAM50 subtypes. Therefore, cNMF provided a more distinctive clustering of immune biology compared with the PAM50 clustering.

Previous immunotherapy studies have demonstrated limited activity in advanced prostate cancer, despite a firmly established role in many malignancies. In many cancers, high PD-L1 expression on immune cells plus high TMB can predict response to checkpoint inhibition.[7,26,27] Furthermore, responses were more likely in prostate cancer patients with high PD-L1 expression.[28] In our cohort, CRPC specimens were more likely to demonstrate lower PD-L1 scores, while TMB and PD-L1 expression were low overall, consistent with previous literature. PD-L1 expression was not significantly different between NMF subtypes, despite significant differences in TMB and gene expression. To date, most immunotherapy trials have focussed on CRPC. However, perhaps the HSPC setting, where higher PD-L1 levels are associated with more aggressive disease,[29] and shorter time to biochemical recurrence,[30] warrants further study. Although
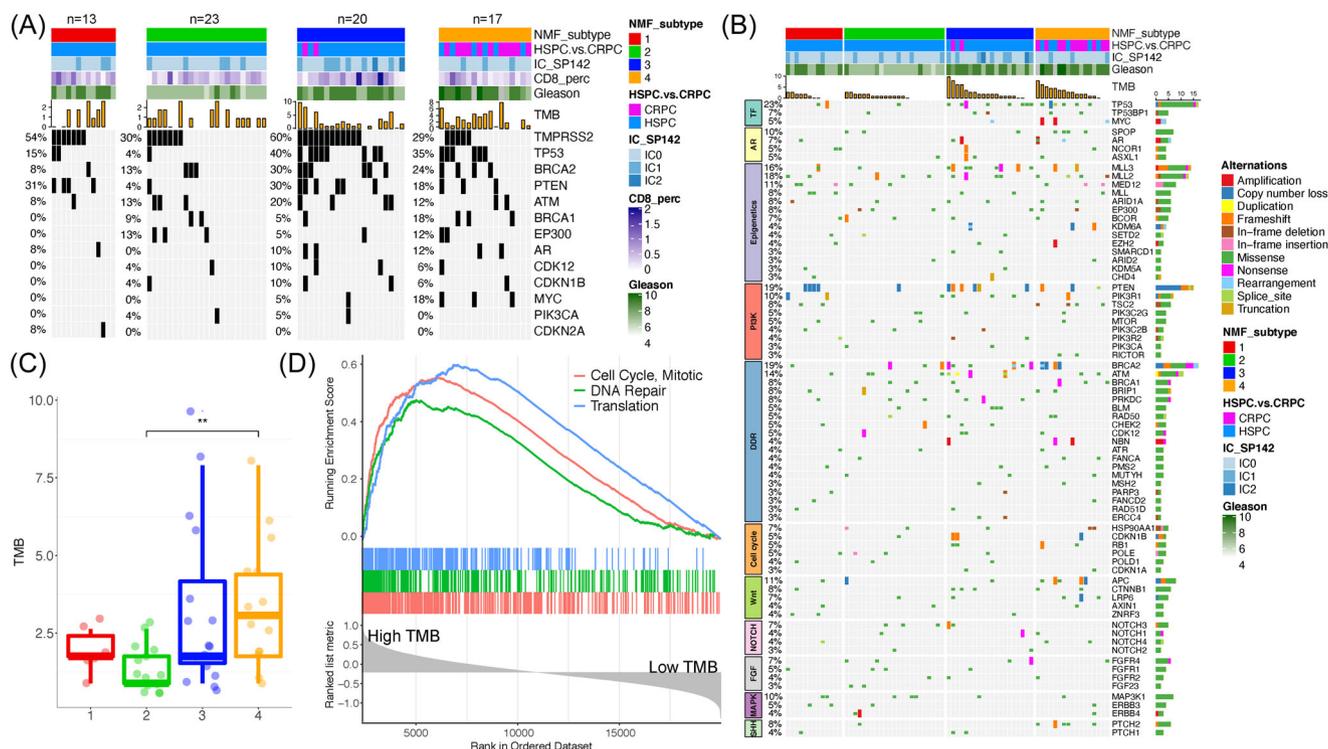
FIGURE 5 Genomic alterations and molecular pathways associated with NMF4 subtypes. (A) Association between NMF subtypes and alterations of some specific genes. (B) Monoprint summarizes altered genes in common molecular cancer signaling pathways/signatures based on NMF subtypes. Genes with higher than 3% alterations are shown. The horizontal bar plots on the right of the oncoprint show the number of samples bearing different types of alterations. (C) The levels of TMB among the NMF subtypes. Statistical significance between NMF subtypes is calculated by Mann–Whitney U test, $**p < 0.01$. (D) Top enriched biological pathways associated with TMB with median cut-off from gene set enrichment analysis (GSEA) using the Reactome database (FDR < 0.05). FDR, false discovery rate; NMF, nonnegative matrix factorization; TMB, tumor mutation burden [Color figure can be viewed at wileyonlinelibrary.com]

higher PD-L1 levels are observed in metastatic specimens within our cohort, the majority of these were obtained in the HSPC setting, a particularly unique aspect of our real-world cohort, supporting the greater potential for immunotherapy in this setting.

In our cohort, NMF4 demonstrated the highest TMB, without corresponding upregulation in immune signatures. Genes related to the mitotic cell cycle, DNA repair, and translation were significantly associated with high TMB. DNA repair defects have been associated with high TMB and improved responses to immunotherapy.[4,31] Although the PD-1 inhibitor pembrolizumab has been approved in all tumors with TMB ≥ 10 mutations/megabase following results of Keynote-158,[32] higher TMB did not correlate with response in the prostate cancer-specific study Keynote-199.[33] You et al.[13] also described a subtype with greater cell cycle expression that was associated with the worst prognosis, as seen in NMF4, which was associated with worst overall survival in our cohort. Furthermore, higher TMB is expected in CRPC specimens, due to the emergence and accumulation of genomic mutations later in the disease trajectory, without necessarily harboring intrinsic sensitivity to checkpoint inhibition.

In some tumor types, high CD8 T-cell infiltration and upregulation of immune signatures such as interferon-γ, are linked to checkpoint inhibitor activity.[34] In prostate cancer, however, there

are conflicting data regarding CD8 T-cell densities and their influence on prognosis.[35,36] In our cohort, such inflammatory signature was observed within NMF1, representing 19% of patients. Despite this, there remains limited activity with checkpoint inhibition in prostate cancer. This may be explained by concurrent upregulation of TGFβ in NMF1. Meng et al.[12] similarly reported an "immune-suppressed" subtype by cNMF clustering, characterized by greater immune cells and TGFβ activation. In our cohort, CD8 density was lower in CRPC specimens. Given the higher PD-L1 and CD8 levels demonstrated in HSPC, further study of immunotherapy in this setting is again warranted.

TGFβ-associated genes are associated with immune evasion and resistance to immunotherapy, mediated by cancer-associated fibroblasts and extracellular matrix cytokines.[37] In metastatic urothelial cancer, TGFβ pathway upregulation is associated with nonresponse to PD-L1 inhibitors and preclinical data suggest that combination therapies inhibiting PD-L1 and TGFβ promote robust antitumor immunity.[34] Jiao et al.[38] demonstrated TGFβ upregulation in murine CRPC models and combined PD-1/PD-L1 plus TGFβ inhibition resulted in increased effector T-cell response and decreased tumor burden. Similarly, Meng et al.[12] demonstrated improved recurrence-free survival in the "immune-activated" subgroup, characterized by a rich immune infiltrate, alongside a low TGFβ signature. Our study

supports this novel treatment combination that may be active in the NMF1 subset.

We acknowledge the limitations of our cohort. Given the small sample size, our findings are hypothesis-generating but under-powered to detect significant differences in clinical outcomes based on NMF subtyping. The clinical data may also reflect differences in disease state and heterogenous use of prior therapies between groups. Furthermore, most patients were treated before the availability of novel therapies and the known emergence of genomic changes following these therapies may not have been entirely captured. Additionally, our analyses utilized archival FFPE specimens that span a 15-year period. Given the potential of RNA degradation, any differences we have identified between HSPC and CRPC, where there may be years separating each specimen, should be interpreted with caution. However, given the consistent total reads and total analyzed reads across our study, RNA degradation did not appear to be an issue.

5 | CONCLUSIONS

Our study demonstrates the utility of cNMF subtyping to identify molecularly distinct prostate cancer cohorts in both HSPC and CRPC and across primary and metastatic specimens. Although NMF1 was associated with the most inflammatory signature, it did not achieve overall survival advantage due to lower TMB and high TGFβ signature expression. NMF4 was enriched for cell cycle and DNA-repair pathways and had greater CRPC samples. Further understanding of the association between molecular subtypes, immune signatures, and mutation profiles will guide future individualized therapeutic strategies.

CONFLICTS OF INTEREST
Kobe Yuen—Employed by Genentech with stock ownership of parent company Roche. Ben Tran—Consulting: Amgen, Astellas, AstraZeneca, Bayer, Bristol-Myers Squibb, Ipsen, Iqvia, Janssen-Cilag, Merck Sharp & Dohme, Novartis, Pfizer/EMD Serono, Roche, Sanofi, and Tomar; Research Funding: Amgen, Astellas, AstraZeneca, Bayer, Bristol-Myers Squibb, Ipsen, Janssen-Cilag, Merck Sharp & Dohme, and Pfizer; Honoraria: Amgen, Astellas, Bristol-Myers Squibb, Janssen-Cilag, Sanofi, and Tolmar; Travel: Amgen and Astellas. Habib Hamidi—Employed by Genentech with stock ownership of parent company Roche. Sanjeev Mariathasan—Employment with and stock/ other ownership interests in Genentech. Shahneen Sandhu—Research grant support to the institution for unrelated work; Merck Sharp & Dohme, Merck Serono, Amgen, Pfizer, AstraZeneca, Endocyte and AAA (a Novartis company), Genentech, and Pfizer; Honorarium to the institution for unrelated work (consultant advisor); AstraZeneca, Amgen, Merck Sharp & Dohme, Merck Serono, Roche/Genentech, Bristol-Myers Squibb, and Novartis. Edward

E. Kadel—Employed by Genentech with stock ownership of parent company Roche; Stock ownership currently or previous 2 years of Clinuvel, Epizyme, Gilead, Mannkind, Merck, Organon, SPDR S&P Biotech ETF, and Teledoc. The remaining authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
Data are openly available in a public repository that does not issue DOIs. Data will be made available following pseudonymization to an appropriate repository.

ORCID
Edward E. Kadel   https://orcid.org/0000-0003-3030-6628

REFERENCES
1. Teo MY, Rathkopf DE, Kantoff P. Treatment of advanced prostate cancer. *Annu Rev Med*. 2019;70:479-499.
2. Martin AM, Nirschl TR, Nirschl CJ, et al. Paucity of PD-L1 expression in prostate cancer: innate and adaptive immune resistance. *Prostate Cancer Prostatic Dis*. 2015;18(4):325-332.
3. Strasner A, Karin M. Immune infiltration and prostate cancer. *Front Oncol*. 2015;5:128.
4. Antonarakis ES, Piulats Rodriguez JMM, Gross-Goupil M, et al. Biomarker analysis from the KEYNOTE-199 trial of pembrolizumab in patients (pts) with docetaxel-refractory metastatic castration-resistant prostate cancer (mCRPC). *J Clin Oncol*. 2020;38(15_suppl): 5526.
5. Abida W, Cheng ML, Armenia J, et al. Analysis of the prevalence of microsatellite instability in prostate cancer and response to immune checkpoint blockade. *JAMA Oncol*. 2019;5(4):471-478.
6. Antonarakis ES, Isaacsson Velho P, Fu W, et al. CDK12-altered prostate cancer: clinical features and therapeutic outcomes to standard systemic therapies, poly (ADP-Ribose) polymerase inhibitors, and PD-1 inhibitors. *JCO Precis Oncol*. 2020;4:370-381.
7. Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019;30(1):44-56.
8. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol*. 2009;27(8): 1160-1167.
9. Zhao SG, Chang SL, Erho N, et al. Associations of luminal and basal subtyping of prostate cancer with prognosis and response to androgen deprivation therapy. *JAMA Oncol*. 2017;3(12):1663-1672.
10. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999;401(6755):788-791.
11. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A*. 2004;101(12):4164-4169.
12. Meng J, Zhou Y, Lu X, et al. Immune response drives outcomes in prostate cancer: implications for immunotherapy. *Mol Oncol*. 2020; 15:1358-1375.
13. You S, Knudsen BS, Erho N, et al. Integrated classification of prostate cancer reveals a novel luminal subtype with poor outcome. *Cancer Res*. 2016;76(17):4948-4958.
14. The Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell*. 2015;163(4): 1011-1025.
15. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43): 15545-15550.
16. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18(1):220.

17. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):218.

18. Forbes SA, Bindal N, Bamford S, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*. 2011;39(database issue):D945-D950.

19. Frampton GM, Fichtenholtz A, Otto GA, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol*. 2013;31(11):1023-1031.

20. Robinson D, Van Allen EM, Wu YM, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;162(2):454.

21. Marzec J, et al. The transcriptomic landscape of prostate cancer development and progression: an integrative analysis. *Cancers (Basel)*. 13, 2021:2.

22. Lapointe J, Li C, Higgins JP, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A*. 2004;101(3):811-816.

23. Markert EK, Mizuno H, Vazquez A, Levine AJ. Molecular classification of prostate cancer using curated expression signatures. *Proc Natl Acad Sci U S A*. 2011;108(52):21276-21281.

24. Mattfeldt T, Wolter H, Kemmerling R, Gottfried HW, Kestler HA. Cluster analysis of comparative genomic hybridization (CGH) data using self-organizing maps: application to prostate carcinomas. *Anal Cell Pathol*. 2001;23(1):29-37.

25. Kim PM, Tidor B. Subsystem identification through dimensionality reduction of large-scale gene expression data. *Genome Res*. 2003;13(7):1706-1718.

26. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet*. 2016;387(10031):1909-1920.

27. Taube JM, Klein A, Brahmer JR, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res*. 2014;20(19):5064-5074.

28. Hansen AR, Massard C, Ott PA, et al. Pembrolizumab for advanced prostate adenocarcinoma: findings of the KEYNOTE-028 study. *Ann Oncol*. 2018;29(8):1807-1813.

29. Gevensleben H, Dietrich D, Golletz C, et al. The immune checkpoint regulator PD-L1 is highly expressed in aggressive primary prostate cancer. *Clin Cancer Res*. 2016;22(8):1969-1977.

30. Vicier C, Ravi P, Kwak L, et al. Association between CD8 and PD-L1 expression and outcomes after radical prostatectomy for localized prostate cancer. *Prostate*. 2021;81(1):50-57.

31. Boudadi K, Suzman DL, Anagnostou V, et al. Ipilimumab plus nivolumab and DNA-repair defects in AR-V7-expressing metastatic prostate cancer. *Oncotarget*. 2018;9(47):28561-28571.

32. Marabelle A, Fakih M, Lopez J, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol*. 2020;21(10):1353-1365.

33. Antonarakis ES, Piulats JM, Gross-Goupil M, et al. Pembrolizumab for treatment-refractory metastatic castration-resistant prostate cancer: multicohort, open-label phase II KEYNOTE-199 study. *J Clin Oncol*. 2020;38(5):395-405.

34. Mariathasan S, Turley SJ, Nickles D, et al. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature*. 2018;554(7693):544-548.

35. Zeigler-Johnson C, Morales KH, Lal P, Feldman M. The relationship between obesity, prostate tumor infiltrating lymphocytes and macrophages, and biochemical failure. *PLOS One*. 2016;11(8):e0159109.

36. Yang Y, Attwood K, Bshara W, et al. High intratumoral CD8(+) T-cell infiltration is associated with improved survival in prostate cancer patients undergoing radical prostatectomy. *Prostate*. 2021;81(1):20-28.

37. Yang L, Pang Y, Moses HL. TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol*. 2010;31(6):220-227.

38. Jiao S, Subudhi SK, Aparicio A, et al. Differences in tumor microenvironment dictate T helper lineage polarization and response to immune checkpoint therapy. *Cell*. 2019;179(5):1177-1190.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.