

1 **Dysregulation of the PRUNE2/PCA3 genetic axis in human prostate**
2 **cancer: From experimental discovery to validation in two**
3 **independent patient cohorts**

4
5 **Richard C. Lauer,^{1,2,*} Marc Barry,^{3,*} Tracey L. Smith,^{4,5} Andrew Maltez Thomas,^{6,15,‡} Jin**
6 **Wu,^{1,7} Ruofei Du,⁸ Ji-Hyun Lee,^{9,10} Arpit Rao,¹¹ Andrey S. Dobroff,^{1,12} Marco A. Arap,^{13,14}**
7 **Diana N. Nunes,¹⁵ Israel T. Silva,¹⁶ Emmanuel Dias-Neto,¹⁵ Isan Chen,¹⁷**
8 **Dennis J. McCance,^{1,7} Webster K. Cavenee,¹⁸ Renata Pasqualini,^{4,5,†} and Wadih Arap,^{4,19,†}**
9

10 ¹University of New Mexico Comprehensive Cancer Center, Albuquerque, New Mexico, USA.

11 ²Division of Hematology/Oncology, Department of Internal Medicine, University of New
12 Mexico School of Medicine, Albuquerque, New Mexico, USA.

13 ³Department of Pathology, University of Utah, Salt Lake City, Utah, USA.

14 ⁴Rutgers Cancer Institute of New Jersey, Newark, New Jersey, USA.

15 ⁵Division of Cancer Biology, Department of Radiation Oncology, Rutgers New Jersey Medical
16 School, Newark, New Jersey, USA.

17 ⁶Department of Biochemistry, Institute of Chemistry, University of São Paulo, São Paulo, SP,
18 Brazil.

19 ⁷Department of Pathology, University of New Mexico, Albuquerque, New Mexico, USA.

20 ⁸Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock,
21 Arkansas, USA.

22 ⁹Department of Biostatistics, University of Florida, Gainesville, Florida, USA.

23 ¹⁰Division of Quantitative Sciences, University of Florida Health Cancer Center, Gainesville,
24 Florida, USA.

25 ¹¹Section of Hematology and Oncology, Department of Medicine, Baylor College of Medicine,
26 Houston, Texas, USA.

27 ¹²Division of Molecular Medicine, Department of Medicine, Albuquerque, New Mexico, USA.

28 ¹³Division of Urology, University of São Paulo Medical School, São Paulo, Brazil.

29 ¹⁴Syrian-Lebanese Hospital, São Paulo, São Paulo, Brazil.

30 ¹⁵Laboratory of Medical Genomics, A.C. Camargo Cancer Center, São Paulo, São Paulo, Brazil.

31 ¹⁶Laboratory of Bioinformatics and Computational Biology, A.C. Camargo Cancer Center, São
32 Paulo, São Paulo, Brazil.

33 ¹⁷MBrace Therapeutics, San Diego, California, USA.

34 ¹⁸Ludwig Institute for Cancer Research, University of California-San Diego, La Jolla, California,
35 USA.

36 ¹⁹Division of Hematology/Oncology, Department of Medicine, Rutgers New Jersey Medical
37 School, Newark, New Jersey, USA.

38

39 **Authorship notes:**

40 *RCL and MB contributed equally to this work. Author order was determined based on order of
41 joining the project.

42 †RP and WA jointly supervised the project.

43 ‡AMT current address: Department of Cellular, Computational and Integrative Biology,
44 University of Trento, Trento, Italy.

45

46 **Address correspondence to:** Wadih Arap, Rutgers Cancer Institute of New Jersey, Newark, 205

47 South Orange Avenue, New Jersey 07101, USA. Phone: 973-972-0366; Email:

48 wadih.arap@rutgers.edu

49

50 **Keywords:** Long Noncoding RNA, oncogene, PCA3, prostate cancer, PRUNE2, tumor

51 suppressor gene

52

53 **Abstract**

54 **BACKGROUND:** We have previously shown that the long non-coding (lnc)RNA *prostate*
55 *cancer associated 3 (PCA3; formerly prostate cancer antigen 3)* functions as a trans-dominant
56 negative oncogene by targeting the previously unrecognized prostate cancer suppressor gene
57 *PRUNE2* (a homolog of the *Drosophila prune* gene), thereby forming a functional unit within a
58 unique allelic locus in human cells. Here we investigated the *PCA3/PRUNE2* regulatory axis
59 from early (tumorigenic) to late (biochemical recurrence) genetic events during human prostate
60 cancer progression.

61 **METHODS:** The reciprocal *PCA3* and *PRUNE2* gene expression relationship in paired prostate
62 cancer and adjacent normal prostate was analyzed in two independent retrospective cohorts of
63 clinically-annotated cases post-radical prostatectomy: a single-institution discovery cohort
64 (n=107) and a multi-institution validation cohort (n=497). We compared the tumor gene
65 expression of *PCA3* and *PRUNE2* to their corresponding expression in the normal prostate. We
66 also serially examined clinical/pathological variables including time to disease recurrence.

67 **RESULTS:** We consistently observed increased expression of *PCA3* and decreased expression
68 of *PRUNE2* in prostate cancer compared with the adjacent normal prostate across all tumor
69 grades and stages. However, there was no association between the relative gene expression levels
70 of *PCA3* or *PRUNE2* and time to disease recurrence, independent of tumor grades and stages.

71 **CONCLUSIONS:** We concluded that upregulation of the lncRNA *PCA3* and targeted
72 downregulation of the protein-coding *PRUNE2* gene in prostate cancer could be early (rather
73 than late) molecular events in the progression of human prostate tumorigenesis but are not
74 associated with biochemical recurrence. Further studies of *PCA3/PRUNE2* dysregulation are
75 warranted.

76 **FUNDING:** We received support from the Human Tissue Repository and Tissue Analysis
77 Shared Resource from the Department of Pathology of the University of New Mexico School of
78 Medicine and a pilot award from the University of New Mexico Comprehensive Cancer Center.
79 RP and WA were supported by awards from the Levy-Longenbaugh Donor-Advised Fund and
80 the Prostate Cancer Foundation. EDN reports research fellowship support from the Brazilian
81 National Council for Scientific and Technological Development (CNPq), Brazil, and the
82 Associação Beneficente Alzira Denise Hertzog Silva (ABADHS), Brazil. This work has been
83 funded in part by the NCI Cancer Center Support Grants (CCSG; P30) to the University of New
84 Mexico Comprehensive Cancer Center (CA118100) and the Rutgers Cancer Institute of New
85 Jersey (CA072720).

86

87

88

89 Introduction

90 Prostate cancer is the most common cancer and the second most common cause of cancer death
91 in men¹, and there continues to be a pressing need for new diagnostic and therapeutic approaches
92 for this disease, as well as better prognostic biomarkers to guide treatment. Long non-coding
93 RNA (lncRNA) species are increasingly recognized as having regulatory functions in
94 tumorigenesis, and nucleic acid-based therapeutics are being developed as a promising means of
95 targeting pathogenic lncRNAs². Several lncRNAs have recently been found to associate with
96 prostate cancer, and the best known of these, *prostate cancer associated 3 (PCA3*; formerly
97 *prostate cancer antigen 3*) has been used clinically for many years as the most specific
98 diagnostic biomarker for prostate cancer^{3,4}; however, its prognostic significance remains
99 uncertain. Strikingly, *PCA3* emerged first only in mammals, with further evolution in primates⁵,
100 and, given aspects of the sequence and genomic organization, we have hypothesized that it might
101 have been introduced into the genome by an ancient oncogenic virus⁶. In humans, *PCA3* has an
102 unusual genomic organization, being present in an antisense direction within an intron of the
103 protein-coding gene *PRUNE2*. Somewhat surprisingly for a molecule that is well-established as a
104 Food & Drug Administration (FDA)- and European Medical Agency (EMA)-approved
105 biomarker, relatively little was known about the biological function of *PCA3* until recently.
106 Ferreira et al. showed that *PCA3* is androgen-regulated and that it promotes prostate cancer cell
107 survival⁷. Subsequently, we have established that *PCA3* downregulates the expression of
108 *PRUNE2* in a rather unusual way: at the RNA level by RNA editing mediated via adenosine
109 deaminase RNA specific (ADAR)-family members⁸. We have showed that expressing ectopic
110 *PCA3* or, alternatively, silencing *PRUNE2* induced cell transformation and cell proliferation *in*
111 *vitro*, increased adhesion and migration of prostate cancer cells, and yielded larger tumors in

112 xenograft tumor models. The opposite biological effects were seen with *PCA3* silencing or
113 ectopic *PRUNE2* expression⁸. Preliminary studies of human prostate cancer samples compared to
114 normal prostate showed increased *PCA3* expression, decreased *PRUNE2* expression, and
115 evidence for RNA editing of these genes. Based on these experimental findings, we proposed
116 that there is a functional molecular axis in human prostate cancer in which *PCA3* acts as a
117 transdominant-negative oncogene to downregulate a previously unrecognized tumor suppressor
118 gene, *PRUNE2*⁸.

119
120 Here we propose that this molecular interplay may serve as a translational target for diagnostic
121 and/or therapeutic intervention in human prostate cancer. First, we present additional correlative
122 evidence from two retrospective post-surgical primary prostate cancer cohorts in support of our
123 experimental model of *PCA3* as a dominant-negative oncogene and *PRUNE2* as a tumor
124 suppressor gene, and for their co-regulation in human prostate cancer. Moreover, we examine the
125 dysregulation of the *PCA3/PRUNE2* regulatory axis across tumors of different grades (patterns),
126 stages, and groups^{9,10}. Finally, we assess whether tumor expression levels of *PCA3* and/or
127 *PRUNE2* are prognostic of biochemical disease recurrence after surgery.

128

129

130 **Methods**

131 *Discovery patient cohort.* Based on a power analysis using gene expression data from our prior
132 work⁸, for the UNMCCC single-institution discovery cohort, we searched the archives of the
133 Department of Pathology at the UNM School of Medicine for at least 100 consecutive patients
134 (final cohort size: n=107) who had a radical prostatectomy as the primary treatment for organ-
135 confined prostate cancer between the years 2001 and 2013 and who had the following clinical
136 and pathological attributes: final post-prostatectomy Gleason Score 7 [either Gleason Group 2
137 (3+4) or Gleason Group 3 (4+3)], pathologic stage pT2 or pT3a, negative surgical margins,
138 negative for seminal vesicle invasion, no evidence of local or distant metastasis, and no prior
139 treatment for prostate cancer. The following additional data were retrospectively abstracted from
140 the individual medical records: age at surgery, race, presence of recurrence, type of recurrence
141 (i.e., biochemical, local, metastatic), and disease-free survival time. Biochemical disease
142 recurrence was defined as a detectable serum prostate-specific antigen (PSA) concentration of at
143 least 0.2 ng/mL post-operatively. Lost to follow up was defined as not having been followed up
144 at the UNMCCC after their urological surgery. All included cases had an independent pathologic
145 re-review by a Board-certified pathologist with expertise in urological oncology (MB), with
146 confirmation of diagnosis, Gleason-based analysis (grading, scoring, and grouping), standard
147 TNM staging, and margin status post-resection. A small number of identified cases (<5%) had to
148 be excluded due to the very limited amount of tumor present.

149
150 *Microdissection of tumor and normal prostate (nonneoplastic prostatic glandular tissue) for the*
151 *discovery cohort.* To obtain tumor for RNA analysis, a representative carcinoma-containing
152 formalin-fixed paraffin embedded (FFPE) block was chosen from each case. Contiguous foci of

153 tumor were marked on the glass slide such that the density of tumor cells was at least 75%. The
154 boundary of the corresponding areas on the tumor block was scored with a blade tip, effectively
155 allowing microdissection of tumor in the process of microtome sectioning. Multiple 10 μ m
156 sections were cut, depending on the area of the tumor focus/foci. In 24 (22.4%) of the cases, we
157 also microdissected areas of nonneoplastic prostatic glandular tissue away from tumor in a
158 similar manner, again also aiming for at least 75% epithelial density.

159
160 *Measurement of PRUNE2 and PCA3 gene expression in the discovery cohort by quantitative RT-*
161 *PCR.* Briefly, gene expression for *PCA3* and *PRUNE2* were determined by quantitative reverse
162 transcription polymerase chain reaction (qRT-PCR) by using TaqMan gene expression assays
163 (Thermo Fisher Scientific) with amplicon detection via a LightCycler 96 (Roche Diagnostics).
164 Gene expression was quantified by the relative logarithmic RT-PCR threshold cycles (Δ Ct)
165 between the target genes and housekeeping control genes¹¹. Specifically, total RNA was
166 extracted from the microdissected FFPE sections using the PureLink™ FFPE Total RNA
167 Isolation Kit (Thermo Fisher Scientific, Cat. No. K1560-02). RNA was quantified on a
168 NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific), and the average A260/A280
169 ratio was 1.94 (range 1.88 to 2.07), indicating optimal quality of the RNA extracted for gene
170 expression assays. RNA was then further quantified with the Qubit® RNA HS Assay Kit
171 (Thermo Fisher Scientific, Cat. No. Q32852) on a Qubit 2.0 (Thermo Fisher Scientific) for
172 accurate RNA concentration. RNA integrity was evaluated with the Agilent RNA 6000 Nano kit
173 (Agilent Technologies, Cat. No. 5067-1511) on an Agilent 2100 Bio analyzer (Agilent
174 Technologies). To remove genomic DNA contamination, RNA samples were treated with 2U of

175 DNase I (Thermo Fisher Scientific, Cat. No. 18068-015) per 2µg of total RNA. All procedures
176 were performed according to the manufacturer's standard protocols.

177
178 Reverse Transcription was performed in triplicate in order to create enough cDNA for the entire
179 project. 500ng RNA in each of three tubes was reverse transcribed with the High-Capacity RNA-
180 to-cDNA™ Kit (Thermo Fisher Scientific, Cat. No. 4387406) in a final volume of 20µl,
181 according to the manufacturer's instructions. Reverse transcription was carried out in a Gene
182 Amp PCR System 9700 (Applied Biosystems) at 37°C for 60 minutes and terminated by 95°C for
183 5 minutes. Then three aliquots were combined for the following experiments.

184
185 For the Thermo Fisher Scientific TaqMan gene expression assay experiments, three
186 (Hs00322421_m1, Hs00999960_m1, and Hs01060890_m1) and two (Hs01371939_g1 and
187 Hs03462121_m1) assays were chosen for target genes *PRUNE2* and *PCA3* respectively,
188 (designated PR1, PR2 and PR3, and PC1 and PC2). Three endogenous controls GAPDH
189 (Hs02758991_g1), HPRT1 (Hs02800695_m1), and UBC (Hs01871556_s1) were selected
190 (designated C1, C2 and C3)¹². Each *PRUNE2* assay and *PCA3* assay was labeled with FAM and
191 paired with a VIC labeled endogenous control in a duplex reaction, with separate reactions to
192 include all of the three endogenous controls. Therefore, a total of fifteen duplex gene expression
193 mixes, nine for *PRUNE2* and six for *PCA3*, was required for all specimens (**Fig. 1 - Source Data**
194 **File 1** [tumor] and **Fig. 1 - Source Data File 2** [normal]).

195
196 Each duplex gene expression assay was then performed in triplicate for all specimens following
197 manufacturer's standard protocols, for a total of 45 expression measures for each case. qRT-PCR

198 was performed with the TaqMan Gene Expression Master Mix (Thermo Fisher Scientific, Cat.
199 No. 4369514) using 1µl of each TaqMan target gene assay (20×FAM) and endogenous controls
200 assay (20×VIC), 1µl of cDNA template (equivalent to 25ng RNA input), and 7µl of RNase-free
201 water for a 20µl final reaction mixture. A non-template control was included in every master mix
202 in every 96-format tray. In addition, in order to evaluate inter-plate variation, we also included
203 one RNA sample, in triplicate, in all the 96 format trays. Analysis of these controls indicated that
204 there were no significant batch effects (data not shown). The qRT-PCR product detection was
205 achieved on a LightCycler 96 (Roche Diagnostics). The cycle program was: at 95°C for 10
206 minutes, followed by 40 cycles at 95°C for 15 seconds and at 60°C for 1 minute. Quantification
207 of target and control genes (Cq) in each sample was performed by LightCycler® 96 SW 1.1
208 (Roche Diagnostics).

209
210 *Validation patient cohort.* For the TGCA patient validation cohort (n=497 patients), we first
211 downloaded clinical data along with the expression of the lncRNA *PCA3* and the *PRUNE2* gene
212 (<http://cancergenome.nih.gov>) with the UCSC Xena browser^{13,14}, together with paired
213 nonneoplastic samples in 52 of the cases (10.5%). The following clinical and pathological
214 characteristics were included in the study: age at diagnosis, vital status, tumor Gleason-based
215 analysis (grading, scoring, grouping), pathological stage, status of biochemical recurrence, and
216 time to recurrence. Gene expression was calculated with RSEM^{15,16}. By using the available
217 dataset, we evaluated *PCA3* and *PRUNE2* gene expression values in terms of tumor versus
218 nonneoplastic prostate, biochemical recurrence, pathologic T stage, Gleason analysis (Grade,
219 Score, and Group), age at pathology-proven diagnosis. Because the regulation of *PRUNE2* by
220 *PCA3* occurs at the RNA level by the formation of an RNA hetero-duplex, we also evaluated the

221 ratio of the expression of the two genes in terms of the clinical and pathological variables for
222 each patient of the cohort.

223
224 *Statistics.* Demographic and clinical variables were summarized with descriptive statistics. For
225 the discovery cohort, the mean and median of gene expressions across multiple control genes and
226 assays were summarized, and these were used as measures for gene expression of *PRUNE2* and
227 *PCA3* relative to endogenous housekeeping controls for each case. More detailed methods are
228 described in the supplemental materials.

229
230 Testing for differences of *PCA3* and *PRUNE2* expression between paired tumor and
231 nonneoplastic prostate expression was by the Wilcoxon signed rank test. The Kruskal-Wallis test
232 was used when comparing three or more groups. Assessment for significant differences of gene
233 expression by recurrence status was by Wilcoxon rank sum test. The Kaplan-Meier product limit
234 method with log-rank test was used to explore the relationship between gene expression levels or
235 the ratio and the time to recurrence. Multivariable Cox proportional hazard modeling was used to
236 fit for the association between time to recurrence and expression levels of *PRUNE2* or *PCA3* or
237 their ratio, while controlling for multiple clinical covariates. All statistical analyses were carried
238 out by using the SAS (9.4) or R software package (R 3.4.5), unless otherwise indicated (R and
239 SAS codes are available in the **Source Code File**).

240
241 *Study approval.* For the discovery cohort, there was University of New Mexico Health Sciences
242 Institutional Review Board (IRB) approval (HRRC15-138), and the study was carried out in
243 accordance with the United States Common Rule.

244 **Results**

245 *Discovery Single-Institutional Cohort*

246 In the initial single-institution discovery cohort from the University of New Mexico
247 Comprehensive Cancer Center (UNMCCC), patients with intermediate-risk (Gleason Score 7;
248 corresponding to Gleason Groups 2 and 3) organ-confined prostate cancer (n=107) met the
249 criteria for inclusion in this study (**Table 1**). Briefly, the mean age of the cohort was 63 years old
250 (ranging from 45 to 84 years old); most patients (85%) were non-Hispanic white, but Hispanic
251 (7.5%), American Indian/Native American (2.8%), and African American (2.8%) men were also
252 represented. All patients had final Gleason Score 7 adenocarcinoma after radical prostatectomy,
253 with 86.9% being 3+4=7 (Gleason Group 2) and 13.1% being 4+3=7 (Gleason Group 3). The
254 pathologic stage distribution was as follows: 74.8% were pT2 and 25.2% were pT3a. Nineteen of
255 the patients (17.8%) had biochemical recurrence discovered during follow up, including one with
256 documented local recurrence and one with documented metastases. Five patients (4.7%) were
257 lost to follow up.

258
259 RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR) were
260 successful in all microdissected tumor samples (n=107). In 24 of these cases (22.4%), we
261 extracted RNA from benign prostatic glandular tissue away from tumor (hereafter termed
262 “normal prostate”: qRT-PCR was successful in all cases for *PRUNE2* (n=24, 100%) and in most
263 cases for *PCA3* (n=21, 87.5%). Comparing *PRUNE2* and *PCA3* expression in prostatic
264 adenocarcinoma with expression in normal prostate (all relative to endogenous housekeeping
265 controls), we found consistent trends for both genes in multiple assays, with lower expression of
266 *PRUNE2* in tumor as compared with normal prostate and higher expression of *PCA3* in tumor as

267 compared with normal prostate (**Fig. 1 - Source Data File 3**). These results are summarized in
268 **Figure 1A** and as follows. Relative to controls, *PCA3* expression was significantly higher in
269 prostatic adenocarcinoma (mean: 2.46; standard deviation: 1.28) than normal prostate (-1.99;
270 2.63, [p-value<0.001]). Relative to controls, *PRUNE2* expression was significantly lower in
271 tumor (mean: -1.48; standard deviation: 0.92) than normal prostate (-0.78; 0.4, [p-value<0.001]).
272
273 We next explored the association between biochemical recurrence and tumor expression levels of
274 *PRUNE2*, *PCA3* and the ratio of *PRUNE2* to *PCA3* expression by using several approaches.
275 First, we compared the gene expression values and their ratio by recurrence status. In patients
276 who recurred compared to those who did not, we found no significant difference in mean
277 expression values of *PRUNE2* (-1.6 vs. -1.58; p-value=0.68), *PCA3* (2.98 vs. 2.43; p-
278 value=0.16), or their ratio (-1.61 vs. -1.21, p-value=0.48). The different expression levels by
279 recurrence were not significant (**Figure 1B, Supplemental Figure S1**). Next, for *PRUNE2*
280 expression, *PCA3* expression, and their ratio, we regrouped the cancer cases according to
281 whether the values were greater than (deemed “high”) or less than/equal to (deemed “low”) their
282 respective mean values. By using the Kaplan-Meier product limit methodology and the log-rank
283 test, we found no significant associations between high or low levels and time to recurrence for
284 *PRUNE2* expression (p-value=0.24), *PCA3* expression (p-value=0.22), (**Figure 2 and Tables 2-**
285 **3**), or their ratio (p-value=0.84). As a further assessment of association between gene expression
286 and time to biochemical recurrence, we used Cox proportional hazards modeling and found no
287 significant associations of time to biochemical recurrence with expression of *PRUNE2* (hazard
288 ratio [HR]: 0.97; 95% confidence interval [CI]: 0.58–1.63, [p-value=0.91]) or *PCA3* (HR: 1.21;
289 95% CI: 0.91–1.6, [p-value=0.19]), or their ratio (HR: 0.99; 95% CI: 0.92–1.1, [p-value=0.82]).

290 By multivariable Cox modeling, we did not find that expression of *PRUNE2*, *PCA3* or their ratio
291 added any additional predictive information for recurrence to that provided by clinical or
292 pathological variables, as presented in supplemental results and **Supplemental Table S1**.

293

294 *Validation/Confirmation Multi-Institutional Cohort*

295 For the analysis of a second multi-institutional validation/confirmation and expansion prostate
296 cancer cohort from The Cancer Genome Atlas (TCGA), patient clinical data (**Table 4**) and gene
297 expression data were available on men with prostate cancer (n=497). Gene expression data from
298 nonneoplastic prostatic tissue (hereafter termed “normal prostate”) were also available in 52
299 (10.5%) of the cases. The basis for the cohort has been described previously¹³. Briefly, the cohort
300 comprised men whose ages ranged from 41 to 78 years old, and who had a radical prostatectomy
301 for primary prostate cancer. The distribution of prostate cancer grades was as follows: Gleason
302 Score 3+3=6 (Group 1), 9.0%; Gleason Scores 3+4=7 or 4+3=7 (Groups 2 and 3), 49.7%;
303 Gleason Score 4+4=8 (Group 4), 12.9%; and Gleason Scores 4+5=9 or 5+4=9, 27.6%; and
304 5+5=10 (Group 5), 0.8%. For pathologic stage, most tumors were pT2c (33%), pT3a (32%), or
305 pT3b (27%), with a small fraction being pT4 (2%). Data on disease recurrence were available for
306 429 men (86.3%), with 58 (11.7% [13.5% of those with follow-up information available]) having
307 biochemical recurrence.

308

309 We compared gene expression levels (\log_2 RNA-Seq by Expectation-Maximization
310 [RSEM]) in prostatic adenocarcinoma to those in normal prostate samples in the data set from
311 TCGA (**Figure 3A**): *PCA3* had significantly increased expression in carcinoma [median: 12.4;
312 interquartile range (IQR): 10.3-13.7] as compared with normal prostate [median: 6.9; IQR: 5.2-

313 9.6, [p-value<0.001]), and *PRUNE2* showed simultaneous lower expression in carcinoma
314 (median: 11.4; IQR: 10.7-12.0) versus normal prostate (median: 12.2; IQR: 11.8-12.6, [p-
315 value<0.001]). As depicted in **Figure 3B**, comparing tumor gene expression across different
316 prostate cancer pathologic grades, *PCA3* expression was significantly lower in tumors with
317 Gleason Score greater than 7 (median: 11.6; IQR: 8.4-13.4) than in tumors with Gleason Score 7
318 (median: 12.8; IQR: 11.3-13.8, [p-value<.001]) or less than 7 (median: 12.5; IQR: 11.8-13.7, [p-
319 value=0.01]). *PRUNE2* showed a small decrease in expression in tumors with Gleason Score
320 greater than 7 (median: 11.3; IQR: 10.4-11.9) as compared with tumors with Gleason Score 7
321 (median: 11.5; IQR: 10.8-12.1, [p-value=0.014]) or less than 7 (median: 11.6; IQR: 11.0-12.1,
322 [p-value=.049]). As shown in **Figure 3C**, comparing tumor gene expression across different
323 tumor pathologic stages, *PCA3* expression was higher in pT2 tumors (median: 12.6; IQR: 11.2-
324 13.8) than in tumors that were pT3 (median: 12.2; IQR: 9.7-13.6, [p-value=0.01]) or pT4
325 (median: 12.1; IQR: 9.6-12.7, [p-value=0.61]). There was no significant difference (p-
326 value>0.05) between *PRUNE2* expression levels between the different tumor stages: pT2
327 (median: 11.4; IQR: 10.7-12.1), pT3 (median: 11.3; IQR: 10.6-12.0), and pT4 (median: 11.7;
328 IQR: 10.8-12.1). We also found that the ratio of *PCA3/PRUNE2* expression showed similar
329 associations with Gleason Score and pathologic stage as were seen with *PCA3* expression (data
330 not shown). Overall, despite the differences in gene expression among tumor grades and stages,
331 the median expression of *PCA3* was significantly higher in all tumor grades and stages than the
332 expression of *PCA3* in normal prostate, and, inversely, the median expression of *PRUNE2* in all
333 tumor grades and stages was significantly less than the expression of *PRUNE2* in normal
334 prostate.

335

336 As shown for the discovery cohort, we also evaluated the relationship between *PCA3* and
337 *PRUNE2* expression levels and recurrence status. We found that patients who had biochemical
338 recurrence after prostatectomy had significantly lower tumor expression levels of *PCA3* (median,
339 11.58; IQR, 8.28-13.14) than those who did not recur [12.51; 10.64-13.71, (p-value<0.01);
340 **Figure 3D**]. However, we did not see an association between tumor *PCA3* expression and
341 biochemical recurrence on multivariable Cox proportional hazards modeling when adjusting for
342 tumor grade, stage, and age at diagnosis [HR, 0.96; 95% CI, 0.87-1.04, (p-value=0.36)], as
343 presented in supplemental results and **Supplemental Table S2**. We did not see a significant
344 association between *PRUNE2* expression in those patients that had biochemical recurrence as
345 compared with those patients who did not recur (**Supplemental Figure S2**).

346

347

348 Discussion

349 Here we assessed the tumor and control adjacent normal prostatic glandular tissue expression of
350 the lncRNA *PCA3* and the protein-coding *PRUNE2* gene in two independent retrospective
351 cohorts of patients with primary organ-confined prostate cancer after treatment by radical
352 prostatectomies (**Figure 4**). As compared with normal prostate, we found that prostate cancer
353 showed consistent increased expression of *PCA3* and consistent decreased expression of
354 *PRUNE2* in tumors across a broad range of pathological attributes (i.e., Gleason grades, scores,
355 groups and stages) in both patient cohorts. These findings support the mechanistic role of a
356 tumor-specific molecular axis in which *PCA3* acts as dominant-negative oncogene and *PRUNE2*
357 as a tumor suppressor gene in human prostate cancer and indicate that the interplay between
358 these genes is dysregulated early in prostate cancer.

359
360 Specifically, when we compared *PCA3* expression in the validation cohort from TCGA, although
361 average expression in all grades, stages, and groups was higher than in normal prostate, we found
362 that among tumors there was significantly decreased *PCA3* expression in tumors with higher
363 grades (Gleason score >7) and in higher stages (>pT2), as compared with lower grades, stages,
364 or groups, respectively. These paradoxical findings are consistent with several early studies^{17,18}
365 and in particular with a recent tissue-based study of *PCA3* expression in prostate cancer¹⁹. In that
366 large cohort study, lower levels of tumor *PCA3* in both biopsy and radical prostatectomy
367 specimens were associated with high-grade tumors, and in radical prostatectomy specimens
368 decreased *PCA3* expression was associated with features of higher stages. Based on these results,
369 it has been proposed that *PCA3* might actually represent a differentiation marker in human
370 prostate cancer¹⁹. The finding of decreasing *PCA3* expression with increasing tumor grades and

371 stages in both our study and others is broadly consistent with another previous study²⁰, which
372 found that the class of antisense intronic RNAs was markedly over-represented among the top
373 transcripts associated with tumor differentiation in human prostate cancer. The finding of an
374 inverse association between *PCA3* expression and increasing grades and stages may also relate to
375 links between *PCA3* expression and androgen receptor (AR) signaling and the likelihood of
376 *PCA3* having an important role in the early steps of prostate cancer carcinogenesis, with a
377 reduced role when the disease is more advanced. Indeed, previous work by our own group and
378 by others indicates that *PCA3* is upregulated by AR signaling⁶⁻⁸, and that *PCA3* is also involved
379 in modulating AR signaling^{7,21}. Interestingly, it also has been shown *in vitro* that *PCA3* silencing
380 sensitizes prostate cancer cells to enzalutamide-induced decreased cell growth²². Alshalalfa et al.
381 suggest that because low pretreatment serum testosterone levels are associated with diseases with
382 higher grades and stages, and because of the relationship between AR signaling and *PCA3*
383 expression, therefore lower *PCA3* expression may reflect the lower serum testosterone in these
384 patients¹⁹. Although, we do not have any data on the pretreatment serum concentration of
385 testosterone and other androgens, our tumor *PCA3* expression findings could perhaps be
386 consistent with that interpretation.

387
388 Because prostate cancers, especially Gleason Score 7 (Groups 2 and 3) tumors, are quite frequent
389 (about half of the total cases) and show divergent clinical behavior, there is great interest in
390 developing prognostic biomarkers for risk stratification. Studies on the association of *PCA3*
391 expression levels with outcome and prognosis show conflicting results²³, and unlike this present
392 study, most prior reports are based on urinary *PCA3* expression²⁴⁻²⁶. Our exploration of the
393 validation cohort from TCGA, which comprised a wide spectrum of tumor grades and stages,

394 revealed an association between lower levels of tumor *PCA3* expression and biochemical
395 recurrence; however, this association was not found after taking grade and stage into account.
396 This finding makes sense, as increasing grade and stage are both variables that are associated
397 with lower *PCA3* expression. In their tissue-based cohort, Alshalalfa et al.¹⁹ also found an
398 association between low *PCA3* levels and adverse outcomes, including biochemical recurrence,
399 metastasis and prostate cancer-specific mortality; however, it is not clear whether such findings
400 are independent of clinical and pathological variables (such as Gleason grade, stage, and group),
401 as a multivariable analysis was not reported. For the discovery cohort of patients, we selected
402 organ-confined, intermediate-risk tumors (Gleason groups 2 and 3, with tumor stages pT2 and
403 pT3) where prognostic information might be expected to be most helpful clinically, to test for an
404 association with outcome. We did not see any association between tumor *PCA3* expression and
405 biochemical recurrence.

406
407 *PRUNE2*, a human homolog of the *Drosophila prune* gene, encodes for a protein with BCH,
408 DHHA2, and PPX1 functional domains⁷. The BCH domain can inhibit the Rho family of
409 proteins, small GTPases with roles in cell transformation, migration and metastasis, and cell
410 cycle progression^{5,27}. Evidence is accumulating that *PRUNE2* might act as a tumor suppressor
411 gene. Loss-of-function mutations have been described in several tumor types, including germline
412 and somatic mutations in parathyroid cancer²⁸ and somatic mutations in solid papillary
413 carcinoma²⁹, while high expression of *PRUNE2* protein correlates with favorable prognosis in
414 neuroblastoma³⁰. Others have shown evidence of inactivating *PRUNE2* mutations in Merkel cell
415 carcinoma³¹ and that the restoration of downregulated *PRUNE2* in oral cancer suppresses tumor
416 cell migration³², further supporting the role of *PRUNE2* as a tumor suppressor. In prostate

417 cancer, the evidence is limited and controversial: an early report found that *PRUNE2* expression
418 was upregulated in prostate cancer and metastases in a small number of samples, and was
419 androgen-inducible in prostate cancer cells⁵. However, a subsequent study on a larger number of
420 samples found that *PRUNE2* expression either decreased or did not increase in aggressive
421 prostate cancer, and that *PRUNE2* expression was not androgen-inducible¹⁷.

422
423 Altogether, the findings in the current study provide additional support for our previous findings⁸
424 that *PRUNE2* acts as a functional tumor suppressor gene in human prostate cancer. Here we
425 described consistently lower expression of *PRUNE2* in prostate cancers of all grades and stages
426 as compared to normal prostate. The findings in our present study are also consistent with the
427 negative regulation of *PRUNE2* by *PCA3* in prostate cancer. We found no significant differences
428 in *PRUNE2* expression across tumor stage, and only a small decrease in expression with
429 increasing tumor grade, suggesting that loss of *PRUNE2* tumor suppressor activity is an early
430 molecular event in prostate cancer. We are not aware of any prior reports of the prognostic
431 significance of tumor *PRUNE2* expression in prostate cancer but, at least in this retrospective
432 study of two independent prostate cancer patient cohorts, we did not find any association
433 between *PRUNE2* expression and biochemical outcomes.

434
435 Strengths of this study include that broadly consistent findings were described in the two
436 independent well-characterized clinically annotated primary prostate cancer cohorts used for
437 analysis, and that the findings were robust across multiple assays in the discovery patient cohort
438 and between the different methods of measurement of gene expression used in the two cohorts.
439 The assessment of *PCA3* expression directly and specifically in tissue (as opposed to urine) is a

440 novelty and a strength as our primary goal was the study of the *PRUNE2/PCA3* regulatory axis
441 in human prostate cancer. We reasoned that the study of tissue expression is likely more
442 informative of tumor biology than traditional urinalysis, not least of all because urinary
443 expression, though very well characterized, could be subject to potential confounding issues such
444 as RNA stability in urine or the contribution of differential urinary shedding. However, from the
445 standpoint of assessment of prognostic information, a drawback of analyzing tissue *PCA3*
446 expression is that the results are not directly comparable to the multiple previous studies that
447 measured urinary *PCA3* scores and ultimately led to FDA and EMA approval for clinical
448 applications in the US and EU. Moreover, while we did find consistent findings with a large
449 tissue cohort study relating *PCA3* expression and biochemical recurrence¹⁹, the analysis
450 presented here was limited in its ability to unequivocally determine the prognostic value of *PCA3*
451 and *PRUNE2* expression as the overall proportion of patients with biochemical recurrences was
452 relatively low. Finally, we were not able to fully address the relationship of reciprocal gene
453 expression of *PCA3* and *PRUNE2* to the outcomes of metastases and prostate cancer-specific
454 deaths, again due to the relative paucity of these events.

455
456 In conclusion, we found consistent upregulation of *PCA3* and downregulation of *PRUNE2* in
457 prostate cancer as compared with normal prostate in two retrospective and independent patient
458 cohorts (summarized in **Figure 4**), supporting that *PCA3* and *PRUNE3* function as an oncogene
459 and a tumor suppressor gene, respectively, in human prostate cancer. The inverse correlation of
460 *PCA3* and *PRUNE2* expression is consistent with our prior findings of a functional interplay
461 between the two genes as part of a unique regulatory unit functioning at a single genetic locus in
462 prostate cancer cells with *PCA3* negatively down-regulating *PRUNE2* expression⁸. The

463 mechanistic dysregulation of *PCA3* and *PRUNE2* is observed across the spectrum of tumor
464 grades and stages, suggesting that this is an early and stable molecular event in prostate cancer.
465 On the other hand, we have not detected any regulatory effects of *PRUNE2/PCA3* in late genetic
466 events such as prostate cancer progressing to biochemical recurrence, which includes the
467 development of local tumor recurrence and/or the development of metastatic disease. The
468 findings presented here represent additional evidence for the functional reciprocal co-regulation
469 of *PCA3* and *PRUNE2* in the setting of early tumorigenesis but not in late events in human
470 prostate cancer. Taken together along with the well-documented specificity of *PCA3*
471 overexpression, our findings establish the *PCA3/PRUNE2* regulatory axis as an attractive early
472 molecular target candidate for intervention in the therapy of human prostate cancer.
473
474

475

476 **Author contributions**

477 RCL, MB, EDN, DJM, RP, WA conceptualized the project; RCL, MB, TLS, AMT, JW, RD,
478 JHL, AR, ASD, MAA, DNN, ITS, EDN, DJM collected and curated the data; RCL, MB, TLS,
479 AMT, JW, RD, JHL, AR, ASD, MAA, DNN, ITS, EDN, DJM performed experiments and
480 provided reagents or critical insights; RCL, MB, TLS, AMT, JW, RD, JHL, AR, ASD, MAA,
481 DNN, ITS, EDN, IC, DJM, WKC,RP, WA analyzed the data; RCL, MB, EDN, DJM, RP, WA
482 wrote the first manuscript draft to which all authors contributed edits; RCL, EDN, DJM, RP, WA
483 obtained funding for the study. RCL, MB, RP, WA, supervised the study.

484

485 **Acknowledgements**

486 We received support from the Human Tissue Repository and Tissue Analysis Shared Resource
487 from the Department of Pathology of the University of New Mexico School of Medicine and a
488 pilot award from the University of New Mexico Comprehensive Cancer Center. RP and WA
489 were supported by awards from the Levy-Longenbaugh Donor-Advised Fund and the Prostate
490 Cancer Foundation. EDN reports research fellowship support from the Brazilian National
491 Council for Scientific and Technological Development (CNPq), Brazil, and the Associação
492 Beneficente Alzira Denise Hertzog Silva (ABADHS), Brazil. This work has been funded in part
493 by the NCI Cancer Center Support Grants (CCSG; P30) to the University of New Mexico
494 Comprehensive Cancer Center (CA118100) and the Rutgers Cancer Institute of New Jersey
495 (CA072720). The results shown here are in part based upon data generated by the TCGA
496 Research Network: <https://www.cancer.gov/tcga>. These data have been presented in part at the
497 2017 American Society of Clinical Oncology Annual Meeting:

498

499 Richard C. Lauer, Marc Barry, Jin Wu, Ji-Hyun Lee, Dennis J. McCance, Ruofei Du, Arpit
500 Rao, Andrey S. Dobroff, Emmanuel Dias-Neto, Isan Chen, Renata Pasqualini, and Wadih Arap.
501 Evaluation of *PRUNE2* and *PCA3* expression as metastasis predictors in Gleason 7 prostate
502 cancer. *J Clin Oncol* 2017; 35:15_suppl, e16583.

503

504 Richard C. Lauer, Marc Barry, Jin Wu, Ji-Hyun Lee, Dennis J. McCance, Ruofei Du, Arpit
505 Rao, Andrey S. Dobroff, Emmanuel Dias-Neto, Isan Chen, Renata Pasqualini, and Wadih Arap.
506 *PRUNE2* and *PCA3* expression in paired non-malignant and tumor specimens from radical
507 prostatectomy patients with Gleason score 7 prostate cancer. *J Clin Oncol* 2017;
508 35:15_suppl, e16582.

509

510 Emmanuel Dias-Neto, Diana N. Nunes, Andrew M. Thomas, Tracey L. Smith, Arpit Rao,
511 Richard C. Lauer, Isan Chen, Wadih Arap, Renata Pasqualini. *PCA3* upregulation in prostate
512 cancer: Analysis in a cohort of 497 subjects from TCGA. *J Clin Oncol* 2017 35:15_suppl,
513 e16578.

514

515 **Competing interest:** The University of New Mexico filed patent applications on *PRUNE2*-
516 related technology (inventors: DNN, EDN, RP, and WA). Those applications were briefly
517 optioned by MBrace Therapeutics, but the applications have since been abandoned and the
518 agreements terminated. WA and RP are founders and equity stockholders of PhageNova Bio,
519 Inc. and of MBrace Therapeutics, Inc.; RP also serves as the Chief Scientific Officer and a paid
520 consultant of PhageNova Bio and as a member of the Board of Directors and a paid consultant

521 for MBrace Therapeutics. IC serves as the Chief Executive Officer of MBrace Therapeutics.
522 Mbrace did not provide financial support for the present work. WKC is a founder and
523 shareholder of Interleukin Combinatorial Therapies, Inc., InVaMet, Inc., and io9, LLC; none of
524 these companies provided funds or participated in the present work. These arrangements are
525 managed in accordance with the established institutional conflict of interest policies for the
526 respective institutions. The remaining authors have declared no potential competing interests
527 exist.

528

529

530 **References:**

- 531 1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin.*
532 Jan 2021;71(1):7-33. doi:10.3322/caac.21654
- 533 2. Arun G, Diermeier SD, Spector DL. Therapeutic Targeting of Long Non-Coding RNAs
534 in Cancer. *Trends Mol Med.* Mar 2018;24(3):257-277. doi:10.1016/j.molmed.2018.01.001
- 535 3. Bussemakers MJ, van Bokhoven A, Verhaegh GW, et al. DD3: a new prostate-specific
536 gene, highly overexpressed in prostate cancer. *Cancer Res.* Dec 1 1999;59(23):5975-9.
- 537 4. de Kok JB, Verhaegh GW, Roelofs RW, et al. DD3(PCA3), a very sensitive and specific
538 marker to detect prostate tumors. *Cancer Res.* May 1 2002;62(9):2695-8.
- 539 5. Clarke RA, Zhao Z, Guo AY, et al. New genomic structure for prostate cancer specific
540 gene PCA3 within BMCC1: implications for prostate cancer detection and progression. *PLoS*
541 *One.* 2009;4(3):e4995. doi:10.1371/journal.pone.0004995
- 542 6. Teixeira AA, Marchio S, Dias-Neto E, et al. Going viral? Linking the etiology of human
543 prostate cancer to the PCA3 long noncoding RNA and oncogenic viruses. *EMBO Mol Med.* Oct
544 2017;9(10):1327-1330. doi:10.15252/emmm.201708072
- 545 7. Ferreira LB, Palumbo A, de Mello KD, et al. PCA3 noncoding RNA is involved in the
546 control of prostate-cancer cell survival and modulates androgen receptor signaling. *BMC Cancer.*
547 Nov 6 2012;12:507. doi:10.1186/1471-2407-12-507
- 548 8. Salameh A, Lee AK, Cardo-Vila M, et al. PRUNE2 is a human prostate cancer
549 suppressor regulated by the intronic long noncoding RNA PCA3. *Proc Natl Acad Sci U S A.* Jul
550 7 2015;112(27):8403-8. doi:10.1073/pnas.1507882112
- 551 9. Gordetsky J, Epstein J. Grading of prostatic adenocarcinoma: current state and prognostic
552 implications. *Diagn Pathol.* Mar 9 2016;11:25. doi:10.1186/s13000-016-0478-2

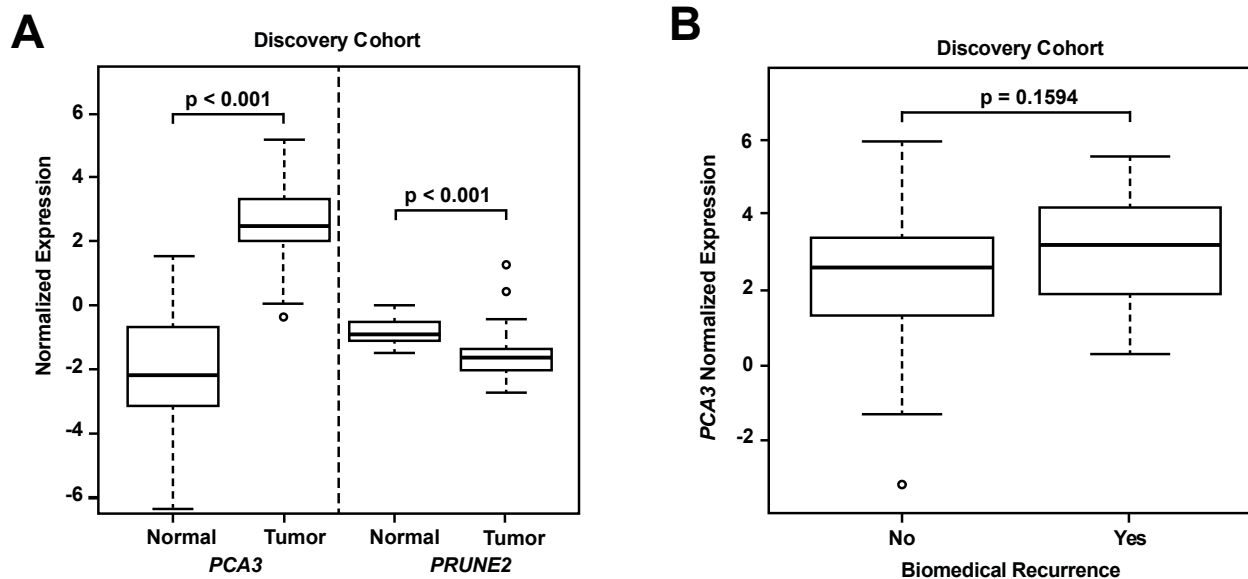
- 553 10. van Leenders G, van der Kwast TH, Grignon DJ, et al. The 2019 International Society of
554 Urological Pathology (ISUP) Consensus Conference on Grading of Prostatic Carcinoma. *Am J*
555 *Surg Pathol*. Aug 2020;44(8):e87-e99. doi:10.1097/PAS.0000000000001497
- 556 11. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time
557 quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods*. Dec 2001;25(4):402-8.
558 doi:10.1006/meth.2001.1262
- 559 12. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time
560 quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome*
561 *Biol*. Jun 18 2002;3(7):RESEARCH0034. doi:10.1186/gb-2002-3-7-research0034
- 562 13. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate
563 Cancer. *Cell*. Nov 5 2015;163(4):1011-25. doi:10.1016/j.cell.2015.10.025
- 564 14. University of North Carolina TCGA Genome Characterization Center. TCGA prostate
565 adenocarcinoma (PRAD) gene expression by RNAseq (polyA+ IlluminaHiSeq). 2017.
566 [https://xenabrowser.net/datapages/?dataset=TCGA.PRAD.sampleMap%2FHiSeqV2&host=https](https://xenabrowser.net/datapages/?dataset=TCGA.PRAD.sampleMap%2FHiSeqV2&host=https%3A%2F%2Ftcga.xenahubs.net&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443)
567 [%3A%2F%2Ftcga.xenahubs.net&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3](https://xenabrowser.net/datapages/?dataset=TCGA.PRAD.sampleMap%2FHiSeqV2&host=https%3A%2F%2Ftcga.xenahubs.net&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443)
568 A443
- 569 15. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or
570 without a reference genome. *BMC Bioinformatics*. Aug 4 2011;12:323. doi:10.1186/1471-2105-
571 12-323
- 572 16. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data
573 via the Xena platform. *Nat Biotechnol*. 06 2020;38(6):675-678. doi:10.1038/s41587-020-0546-8

- 574 17. Salagierski M, Verhaegh GW, Jannink SA, Smit FP, Hessels D, Schalken JA. Differential
575 expression of PCA3 and its overlapping PRUNE2 transcript in prostate cancer. *Prostate*. Jan 1
576 2010;70(1):70-8. doi:10.1002/pros.21040
- 577 18. Balcerczak E, Mirowski M, Sasor A, Wierzbicki R. Expression of p65, DD3 and c-erbB2
578 genes in prostate cancer. *Neoplasma*. 2003;50(2):97-101.
- 579 19. Alshalalfa M, Verhaegh GW, Gibb EA, et al. Low PCA3 expression is a marker of poor
580 differentiation in localized prostate tumors: exploratory analysis from 12,076 patients.
581 *Oncotarget*. Aug 1 2017;8(31):50804-50813. doi:10.18632/oncotarget.15133
- 582 20. Reis EM, Nakaya HI, Louro R, et al. Antisense intronic non-coding RNA levels correlate
583 to the degree of tumor differentiation in prostate cancer. *Oncogene*. Aug 26 2004;23(39):6684-
584 92. doi:10.1038/sj.onc.1207880
- 585 21. Lemos AE, Ferreira LB, Batoreu NM, de Freitas PP, Bonamino MH, Gimba ER. PCA3
586 long noncoding RNA modulates the expression of key cancer-related genes in LNCaP prostate
587 cancer cells. *Tumour Biol*. Aug 2016;37(8):11339-48. doi:10.1007/s13277-016-5012-3
- 588 22. Ozgur E, Celik AI, Darendeliler E, Gezer U. PCA3 Silencing Sensitizes Prostate Cancer
589 Cells to Enzalutamide-mediated Androgen Receptor Blockade. *Anticancer Res*. Jul
590 2017;37(7):3631-3637. doi:10.21873/anticancer.11733
- 591 23. Loeb S, Partin AW. Review of the literature: PCA3 for prostate cancer risk assessment
592 and prognostication. *Rev Urol*. 2011;13(4):e191-5.
- 593 24. Loeb S, Bruinsma SM, Nicholson J, et al. Active surveillance for prostate cancer: a
594 systematic review of clinicopathologic variables and biomarkers for risk stratification. *Eur Urol*.
595 Apr 2015;67(4):619-26. doi:10.1016/j.eururo.2014.10.010

- 596 25. Lemos AEG, Matos ADR, Ferreira LB, Gimba ERP. The long non-coding RNA PCA3:
597 an update of its functions and clinical applications as a biomarker in prostate cancer. *Oncotarget*.
598 Nov 12 2019;10(61):6589-6603. doi:10.18632/oncotarget.27284
- 599 26. Fenstermaker M, Mendhiratta N, Bjurlin MA, et al. Risk Stratification by Urinary
600 Prostate Cancer Gene 3 Testing Before Magnetic Resonance Imaging-Ultrasound Fusion-
601 targeted Prostate Biopsy Among Men With No History of Biopsy. *Urology*. Jan 2017;99:174-
602 179. doi:10.1016/j.urology.2016.08.022
- 603 27. Iwama E, Tsuchimoto D, Iyama T, et al. Cancer-related PRUNE2 protein is associated
604 with nucleotides and is highly expressed in mature nerve tissues. *J Mol Neurosci*. Jun
605 2011;44(2):103-14. doi:10.1007/s12031-010-9490-2
- 606 28. Yu W, McPherson JR, Stevenson M, et al. Whole-exome sequencing studies of
607 parathyroid carcinomas reveal novel PRUNE2 mutations, distinctive mutational spectra related
608 to APOBEC-catalyzed DNA mutagenesis and mutational enrichment in kinases associated with
609 cell migration and invasion. *J Clin Endocrinol Metab*. Feb 2015;100(2):E360-4.
610 doi:10.1210/jc.2014-3238
- 611 29. Alsadoun N, MacGrogan G, Truntzer C, et al. Solid papillary carcinoma with reverse
612 polarity of the breast harbors specific morphologic, immunohistochemical and molecular profile
613 in comparison with other benign or malignant papillary lesions of the breast: a comparative study
614 of 9 additional cases. *Mod Pathol*. Sep 2018;31(9):1367-1380. doi:10.1038/s41379-018-0047-1
- 615 30. Machida T, Fujita T, Ooo ML, et al. Increased expression of proapoptotic BMCC1, a
616 novel gene with the BNIP2 and Cdc42GAP homology (BCH) domain, is associated with
617 favorable prognosis in human neuroblastomas. *Oncogene*. Mar 23 2006;25(13):1931-42.
618 doi:10.1038/sj.onc.1209225

- 619 31. Harms PW, Vats P, Verhaegen ME, et al. The Distinctive Mutational Spectra of
620 Polyomavirus-Negative Merkel Cell Carcinoma. *Cancer Res.* Sep 15 2015;75(18):3720-3727.
621 doi:10.1158/0008-5472.CAN-15-0702
- 622 32. Su SC, Yeh CM, Lin CW, et al. A novel melatonin-regulated lncRNA suppresses TPA-
623 induced oral cancer cell motility through replenishing PRUNE2 expression. *J Pineal Res.* Oct
624 2021;71(3):e12760. doi:10.1111/jpi.12760
- 625
- 626

627 **Figures:**



628

629

630 **Figure 1: Analyses of discovery prostate cancer cohort. (A)** *PRUNE2* and *PCA3* expression in

631 tumor (**Fig. 1 - Source Data File 1**) and nonneoplastic (**Fig. 1 - Source Data File 2**) prostatic

632 samples. Calculated values available in **Fig. 1 - Source Data File 3**. **(B)** Tumor *PCA3* expression

633 by biochemical recurrence status. *PCA3* expression in patients without versus with biochemical

634 recurrence in the discovery cohort. No significant difference in median expression was seen in

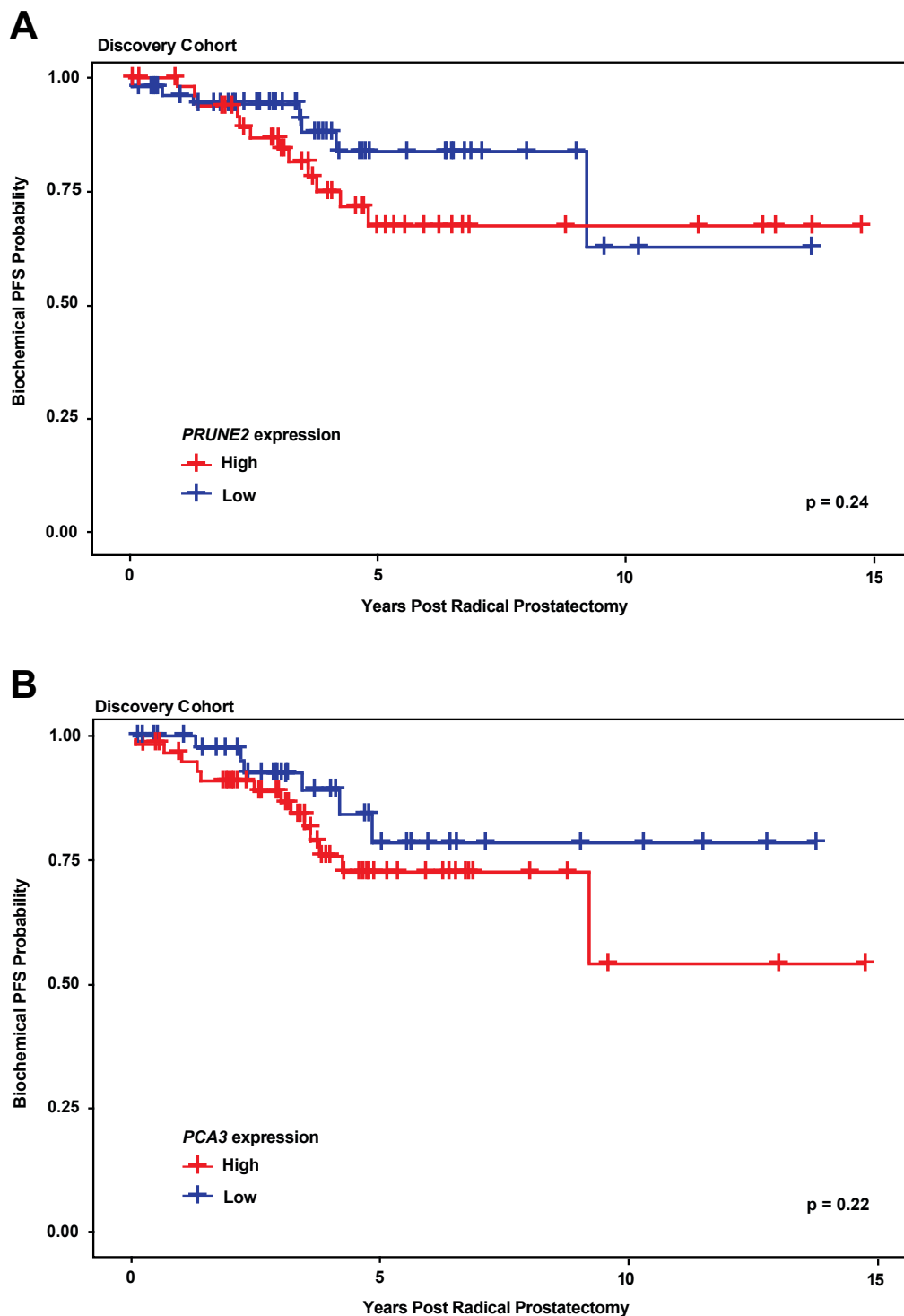
635 this cohort. Box plots of gene expression (normalized expression) in the discovery cohort. The

636 horizontal line within each box represents the median value, while the box represents the

637 interquartile range, and the whiskers extend out to 1.5 times the interquartile range. Outliers are

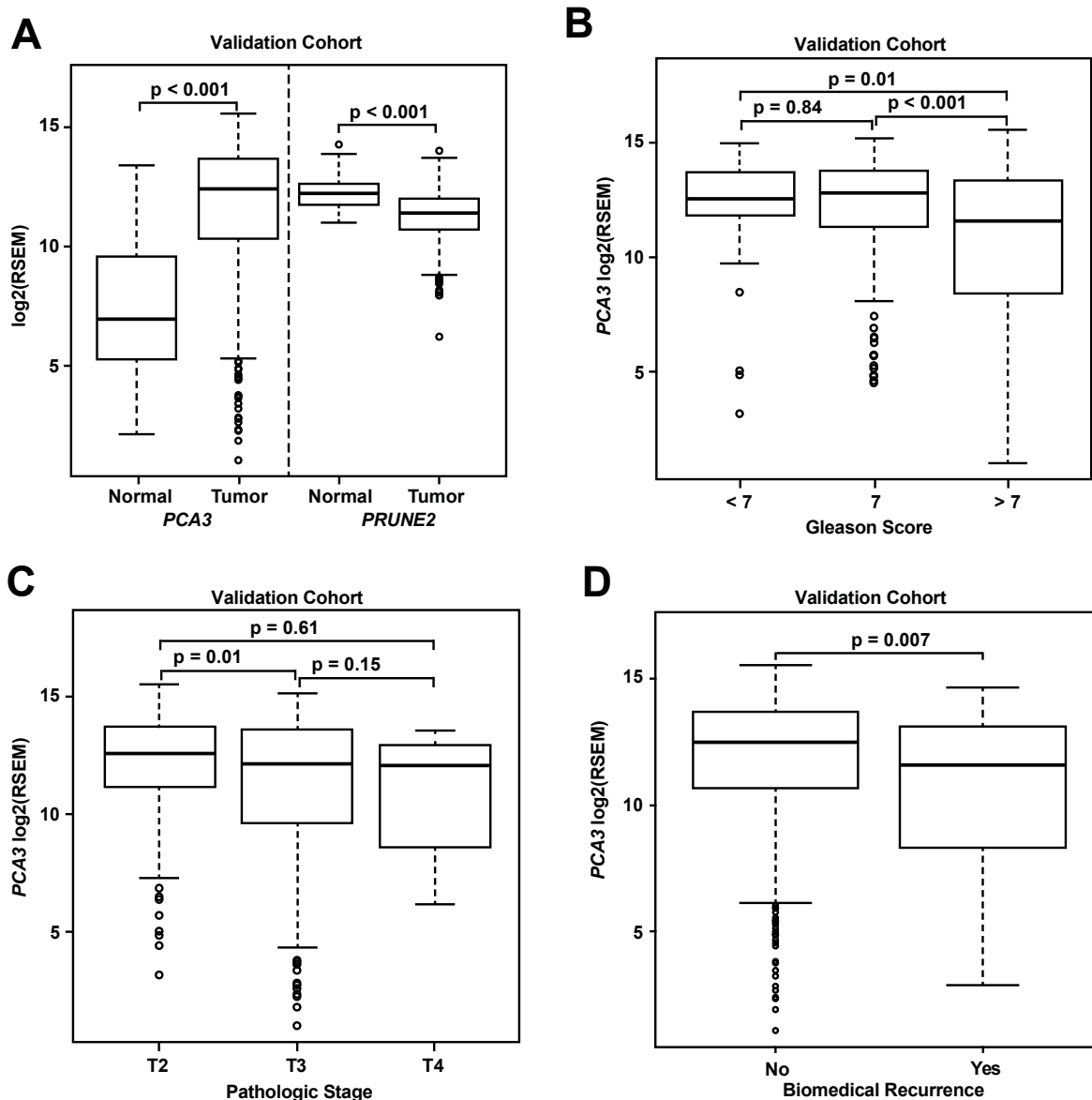
638 represented by circles. P-values are noted for the indicated comparisons.

639



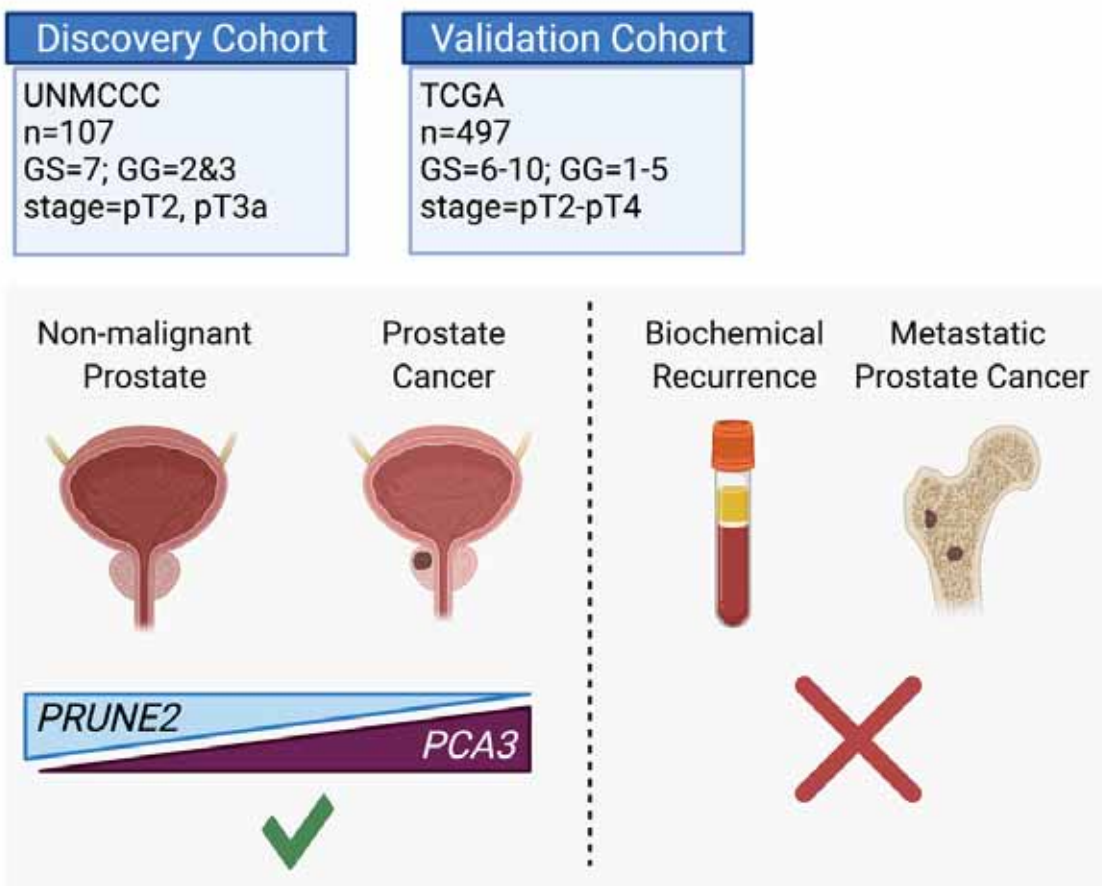
640

641 **Figure 2.** Kaplan-Meier curves illustrating time to event (biochemical recurrence) among
642 patients post-prostatectomy from the discovery cohort, stratified by “high” gene expression
643 (greater than mean expression, red line) versus “low” gene expression (less than or equal to mean
644 expression, blue line), for (A) tumor *PRUNE2* expression, and (B) tumor *PCA3* expression.
645 There is was no significant association of high versus low expression levels and time to
646 recurrence by log rank testing for either *PRUNE2* or *PCA3*.



647

648 **Figure 3.** Analyses of TCGA prostate cancer validation cohort. (A) *PCA3* and *PRUNE2* expression in
 649 nonneoplastic prostatic glandular tissue and in prostatic adenocarcinoma in the TCGA cohort illustrating
 650 consistent gene expression differences between tumor and nonneoplastic prostate in both cohorts. (B, C)
 651 *PCA3* expression in the cohort from TCGA across Gleason grades (B) showing lower expression in
 652 higher grade (>7) tumors and across tumor stages (C) showing lower expression in higher stage tumors.
 653 All illustrated tumor grades and stages show higher expression than nonneoplastic prostatic epithelium.
 654 (D) *PCA3* expression in patients without versus with biochemical recurrence in the TCGA cohort. In the
 655 TCGA cohort, lower *PCA3* median expression was associated with biochemical recurrence. Box plots of
 656 gene expression in the TCGA cohort is reported as \log_2 RSEM data. The horizontal line within each box
 657 represents the median value, while the box represents the interquartile range, and the whiskers extend out
 658 to 1.5 times the interquartile range. Outliers are represented by circles. P-values are noted for the
 659 indicated comparisons. (RSEM: RNA-Seq by Expectation-Maximization).
 660



661
662 **Figure 4. Graphical summary of the analyses.** *PCA3* and *PRUNE2* gene expression
663 relationship in paired prostate cancer and adjacent normal prostate was analyzed in two
664 independent retrospective cohorts of clinically-annotated cases post-radical prostatectomy: a
665 single-institution discovery cohort (n=107) and a multi-institution validation cohort (n=497). We
666 also serially examined clinical/pathological variables including time to disease recurrence.
667 Created with BioRender.com
668

669 **Tables:**

670 **Table 1.** Discovery Cohort: clinicopathologic features of the 107 patients in this study who had
 671 radical prostatectomy at UNM for localized prostatic adenocarcinoma (Further details in **Table 1**
 672 – **Source Data File 1** and **Table 1 – Supplemental Figure 1**)

Variable	Number	% or Mean (STD; Range)
Age (years)	107	62.8 (8.4; 45-84)
Race		
Non-Hispanic White	91	85
Hispanic	8	7.5
American Indian	3	2.8
African American	3	2.8
Other	2	1.9
Post-prostatectomy Gleason Grade (Grade Group)		
3 + 4 = 7 (Grade Group 2)	93	86.9
4 + 3 = 7 (Grade Group 3)	14	13.1
Pathologic Stage		
pT2	80	74.8
pT3a	27	25.2
Biochemical Recurrence Status		
No	83	77.6
Yes	19	17.8
LTF	5	4.7

673 Abbreviations: STD, standard deviation; LTF, lost to follow up

674

675

676

677 **Table 2.** Number of patients at risk over time (see **Figure 2A**)

678

	Years Post Radical Prostatectomy				679
<i>PRUNE2</i> Expression	0	5	10	15	
High	51	15	5	0	680
Low	56	14	2	0	681

682 **Table 3.** Number of patients at risk over time (see **Figure 2B**)

683

	Years Post Radical Prostatectomy				684
<i>PCA3</i> Expression	0	5	10	15	
High	59	16	2	0	
Low	48	13	5	0	

685 **Table 4.** Validation cohort: clinicopathologic features of the 497 patients in the prostate cancer
 686 TCGA data set, with a total of 549 tissue samples included.

Variable	Number	% or Mean (STD; Range)
Age (years)	497	61 (6.8; 41-78)
Vital Status		
Alive	487	97.9
Dead	10	2.1
Sample Type		
Primary tumor	497	
Normal (non-malignant) prostate	52	
Post-prostatectomy Gleason Grade (Grade Group)		
6 (Grade Group 1)	45	9
7 (Grade Groups 2 & 3)	247	49.7
8 (Grade Group 4)	64	12.9
9 (Grade Group 5)	137	27.6
10 (Grade Group 5)	4	0.8
Pathologic Stage		
pT2a	13	2.6
pT2b	10	2
pT2c	164	33
pT3a	158	32
pT3b	135	27
pT4	10	2
Unknown	7	1.4
Biochemical Recurrence Status		
No	371	74.6
Yes	58	11.7
Unknown	68	13.7

687 Abbreviation: STD, standard deviation

688

689

690 **Source Data:**

691 **Table 1 - Source Data File 1. Discovery Cohort.** Clinicopathologic features of the 107 patients
692 in this study who had radical prostatectomy at UNM for localized prostatic adenocarcinoma

693
694 **Figure 1 - Source Data File 1. Analyses of discovery prostate cancer cohort.** Raw values of
695 *PRUNE2* and *PCA3* expression in tumor prostatic samples.

696
697 **Figure 1 - Source Data File 2. Analyses of discovery prostate cancer cohort.** Raw values of
698 *PRUNE2* and *PCA3* expression in nonneoplastic prostatic samples.

699
700 **Figure 1 - Source Data File 3. Analyses of discovery prostate cancer cohort.** Calculated
701 values of *PRUNE2* and *PCA3* expression in tumor and nonneoplastic prostatic samples.

702
703 **Source Code File. R code and SAS code of descriptive statistics.**
704

705 **Supplemental Material:**

706 **Results**

707 In the discovery cohort, to further assess the association between time to biochemical recurrence
708 and the gene expression of *PRUNE2* and *PCA3* and the ratio of their expressions, and to take into
709 account possible confounding by clinicopathological variables, we performed multivariate Cox
710 proportional hazards regression modeling (raw data available in Supplemental File 1 tables and
711 code). As there may be possible correlations between tumor Gleason Grade (GG) and tumor
712 pathologic stage (PS), we created a composite categorical variable (“GG-PS”) representing the
713 four possibilities in the discovery cohort: 3+4/pT2, 4+3/pT2, 3+4/pT3a, 4+3/pT3a. We then used
714 two approaches to model outcome.

715
716 In the first approach, we fit four different Cox models. The explanatory variables in these models
717 were as follows: model 1 (“GG_PS model”) – GG_PS only; model 2 (“*PRUNE2* model”) –
718 *PRUNE2* expression, age, race, GG_PS, and interaction between GG_PS and *PRUNE2*
719 expression; model 3 (“*PCA3* model”) – *PCA3* expression, age, race, GG_PS, and interaction
720 between GG_PS and *PCA3* expression; and model 4 (“Ratio model”) – *PRUNE2/PCA3* ratio,
721 age, race, GG_PS, and interaction between GG_PS and the ratio. The goodness-of-fit of the
722 models were compared using the Akaike Information Criterion (AIC), and the results are
723 summarized in **Supplemental Table S1**. As a lower value of AIC indicates a better association
724 fit, the model comparison indicates that the model with GG_PS only (model 1) represents the
725 best fit for the data, and does not suggest that the expression of *PRUNE2*, *PCA3* or their ratio
726 adds to the ability of pathology grade and stage to predict biochemical recurrence.

727

728 As a second multivariable approach to an assessment of the association of time to biochemical
729 recurrence and gene expression, we created a multivariable model, including the following
730 variables: expression of *PRUNE2*, expression of *PCA3*, age, race, the composite variable
731 GG_PS, interaction between *PRUNE2* expression and GG_PS, and interaction between *PCA3*
732 expression and GG_PS. In this case, a stepwise selection algorithm was used for model selection,
733 and only the model with Gleason score and pathologic stage (GG_PS) was selected, as none of
734 the other variables had a p-value less than the specified significance level of 0.25 (data not
735 shown).

736
737 In the TCGA cohort, we used Cox proportional hazards modeling to assess the association of
738 tumor *PCA3* expression, adjusting for tumor grade, tumor stage, and age at diagnosis. The results
739 are summarized in **Supplemental Table S2**. Briefly, on multivariable modeling, there was an
740 association between tumor grade and stage with recurrence, we did not find that tumor *PCA3*
741 expression was associated with biochemical recurrence [HR, 0.96; 95% CI 0.87-1.04, (p-
742 value=0.36)].

743

744 **Methods**

745 *Statistical Analysis for Quantifying the expression of PCA3 and PRUNE2.* There were
746 combinations of assays and control genes used for quantifying the expression of *PCA3* and
747 *PRUNE2* in this study. Explicitly, there were 9 duplex mixes for *PRUNE2*: PR1C1, PR1C2,
748 PR1C3, PR2C1, PR2C2, PR2C3, PR3C1, PR3C2, PR3C3; and 6 duplex mixes for *PCA3*:
749 PC1C1, PC1C2, PC1C3, PC2C1, PC2C2, PC3C3, where the first 3 letters denote an assay and
750 last two letters denote a control gene being used in a particular run. For example, PC2C2 denotes

751 the second assay for *PCA3* (Hs03462121_m1, detailed in Methods) and the second endogenous
752 control gene (Hs02800695_m1, detailed in Methods) were used for that specific experiment.

753 C_{T0} is to denote the logarithmic number of PCR cycle when the fluorescent signal passes a
754 threshold value. Let $\Delta C_{T0} = C_{T0} \text{ study gene} - C_{T0} \text{ control gene}$ and we had $-\Delta C_{T0}$ to quantify the
755 gene expression (relative to a control gene), resulting in a positive value meaning an upregulated
756 gene's expression.

757 The experiment was completed 3 times for each gene duplex mix, e.g. we have three data points
758 of PC2C2 measure for a tumor sample. The median of the three $-\Delta C_{T0}$ values is summarized to
759 estimate the gene expression of a particular gene duplex mix. We then looked at both mean and
760 median of 9 estimates for *PRUNE2* and 6 estimates for *PCA3*, separately (data not shown). We
761 did not see any significant difference utilizing mean or median in this or subsequent analyses.

762

763 **Supplemental Tables:**

764 **Supplemental Table S1.** Discovery cohort multivariable model

Model	Akaike Information Criterion Score
PS_GG model	157.2
<i>PRUNE2</i> model	162.9
<i>PCA3</i> model	166.2
Ratio model	162.2

765

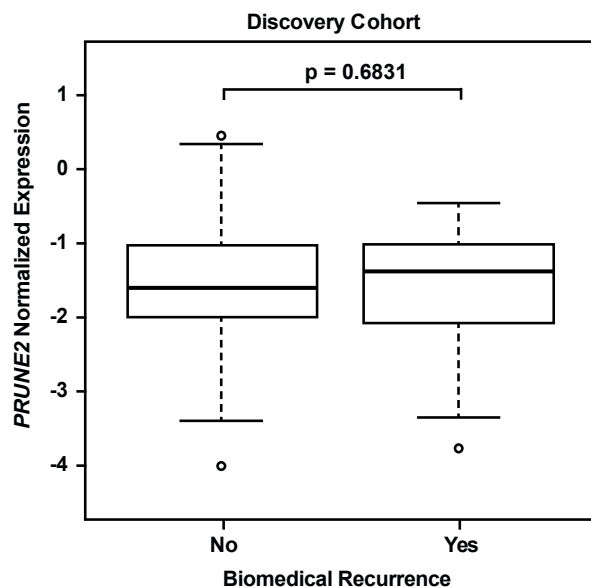
766

767 **Supplemental Table S2.** Validation cohort multivariable cox model

	Hazard Ratio (HR)	95% CI	p-value
<i>PCA3</i>	0.963	0.888 – 1.044	0.36
Gleason Grade	1.558	1.147 – 2.117	<0.01
Pathologic Stage T3	3.596	1.360 – 9.512	<0.01
Pathologic Stage T4	1.860	0.206 – 16.82	0.58
Age at Diagnosis	1.000	0.960 – 1.043	0.99

768

769 **Supplemental Figures:**



770

771 **Supplemental Figure S1:** Discovery cohort - no significant difference in tumor *PRUNE2*
772 expression by biochemical recurrence status.

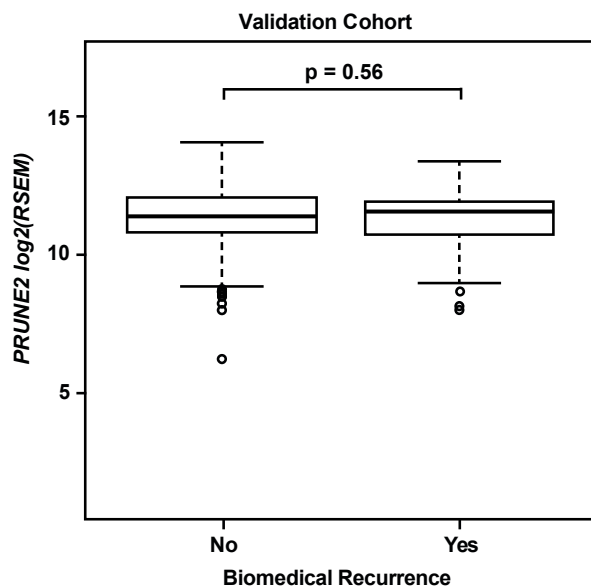
773

774

775

776

777



778

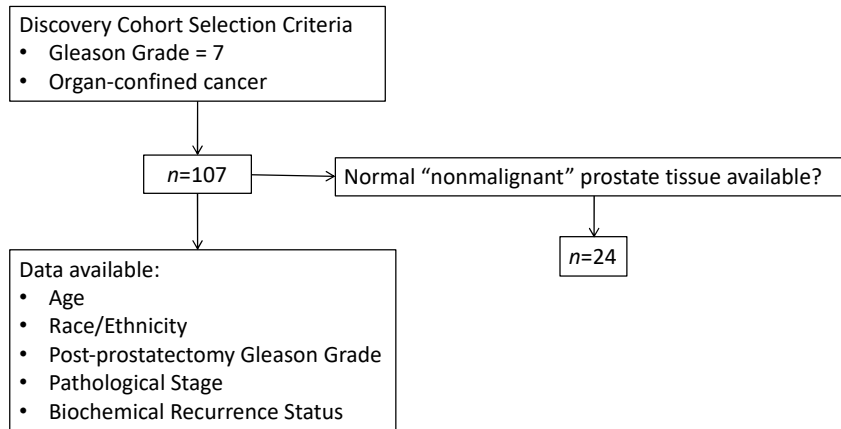
779

780 **Supplemental Figure S2:** Validation Cohort – no significant difference in tumor *PRUNE2*
781 expression by biochemical recurrence status

782

783

784



785

786 **Table 1 - Supplemental Figure 1:** Flow diagram of Discovery Cohort selection criteria and
787 clinicopathological characteristics available.

788